222
27. 8. 70

The British Journal of Educational Psychology

(Incorporating the "Forum of Education")

EDITED BY

P. E. VERNON

AIDED IN THE SELECTION OF PAPERS BY

SIR CYRIL BURT
R. A. C. OLIVER
E. A. PEEL
C. W. VALENTINE
S. WISEMAN
HELEN M. WODEHOUSE

AND THE FOLLOWING REPRESENTATIVES
OF OTHER COUNTRIES :

L. CARMICHAEL
ARNOLD GESELL
I. LORGE
I. MEYERSON
JEAN PIAGET
FRED J. SCHONELL

ASSISTANT EDITOR
L. B. BIRCH

Volume XXIX
1959

BUREAU OF EDL. & PSYL. RESEARCH
LIBRARY

Published by
METHUEN AND CO., LTD., 36, ESSEX STREET, STRAND, LONDON, W.C.2,
for the British Journal of Educational Psychology, Ltd.
Registered Offices : 42, Hill Street, Birmingham, 5.
on behalf of
THE BRITISH PSYCHOLOGICAL SOCIETY and THE ASSOCIATION
OF TEACHERS IN COLLEGES AND DEPARTMENTS OF EDUCATION

All rights reserved.

[illegible]
Date 27.8.70
Acc. No. J 222

VOLUME XXIX. FEBRUARY, 1959 PART I

THE LONG-TERM FOLLOW-UP OF ASSESSMENTS AT AGE ELEVEN

BY JOHN NISBET
(*University of Aberdeen.*)
AND
JIM BUCHAN
(*Culter Secondary School, Aberdeenshire.*)

SUMMARY. 1.—The careers of 102 students who entered Aberdeen University from Aberdeen City schools in 1953, were traced from the end of primary school to fourth year at university.

2.—Of those who entered Aberdeen County secondary schools in 1950, 153 sat the Leaving Certificate at the end of secondary education, and forty-six entered Aberdeen University; their careers were traced from primary school to first-year university.

3.—In both investigations scores in tests at the age of eleven-plus were compared with subsequent performance in Leaving Certificate and university examinations, as a check on the accuracy of the tests in predicting high scholastic achievement six to ten years later.

4.—The results show that even in this highly selected group, the tests at age eleven-plus are not without validity; but the agreement between test scores and university performance is not high, particularly in science courses.

5.—Correlations between five primary school assessments in the City group and fourth-year university performance ranged from +0·05 to +0·28.

6.—In the County group, out of 77 who gained university entrance qualifications, 19 had scored below I.Q. 110 in the test at the end of primary school.

7.—Examples of poor subsequent performance by high scorers in the selection tests at eleven plus, and of excellent performance by low scorers, indicate that, within this group and over this interval of years, accurate prediction of scholastic attainment is not possible.

I.—INTRODUCTION.

SELECTION for secondary education is not principally concerned with the potential university graduate. However, it is of importance that a system of selection should not be such as to exclude possible future graduates from courses which lead to university entrance qualifications. The first of the two follow-up investigations described here was concerned with a group whose formal education continued until graduation from university. Could scholastic achievement in such a group be predicted accurately, or at all, at the age of eleven? An answer to this question was sought by tracing back the school careers of 102 university students who had attended Aberdeen City schools. The second inquiry took the question a stage further: if prediction of scholastic success is only moderately accurate at age eleven, is there any evidence that pupils who obtain low scores in selection tests at that age might ultimately succeed in university courses, given the opportunity through a highly flexible organization of secondary school courses? Results are presented from an inquiry in the County of Aberdeen, in which the careers of pupils entering the secondary schools in 1950 were followed up as far as the first year of university study.

Bureau Ednl. Psy. Research
DAVID HARE TRAINING COLLEGE
Dated

II.—ABERDEEN CITY FOLLOW-UP.

(*a*) *Subjects, selection tests, and criteria of subsequent performance.*—In Octobe 1953, 102 boys and girls from Aberdeen City secondary schools entered the University of Aberdeen; 88 of the 102 sat the transfer examination at the end of primary schooling in 1946, 13 sat it in 1947 and one sat it in 1945. The examination followed the same pattern each year: two Moray House intelligence tests (in 1947 MHT 33 and 34), a Moray House Arithmetic test (in 1947, MHA 15), a Moray House English test (in 1947, MHE 15), and teachers' estimates of primary school attainment scaled on the test totals to allow comparisons between schools. Each year, the scores in each measure are restandardized in Aberdeen to give a mean of 100 and a standard deviation of 15 for each test in the complete 11-year-old age group of the City. For this reason the grouping together of all 102 students was felt to be justified, since these new standardized scores are roughly comparable from test to test and from year to year.

Secondary school performance was assessed on the basis of performance in the Scottish Leaving Certificate, taken usually at age 17. Obviously, this is inadequate as a measure of the extent to which a pupil has profited from secondary education, but it was adopted here as it is the basis for university entrance. There are two grades of pass in each subject, Higher and Lower. The minimum standard for university entrance is three Higher passes and two Lower passes, or four Higher passes (in certain specified combinations of subjects). To permit a scale of performance in this examination, two Lower passes were counted as equivalent to one Higher pass.

Finally, there was university performance which was assessed at three stages, at the end of the first, third and fourth years. First year performance was assessed on the basis of merit certificates gained in first-year subjects—two points for a first-class certificate and one for a second-class. In Scotland the majority of students in Arts and Science take Ordinary pass degrees; and most of the Ordinary graduates in the group completed their courses in 1956. Most of the Honours students graduated in 1957. The medical students have not yet graduated, but completed the first part of their Third Professional Examination in 1957. A five point scale was, therefore, constructed. (The following details apply to fourth-year performance: third year ratings were on similar lines). In Arts and Science, A ratings were given to first-class Honours graduates, and B ratings to second-class Honours. C ratings were third-class Honours, or Ordinary M.A. or B.Sc. graduates with a good record of successes. Rating D indicated a poor Ordinary pass after failures; and E failure to complete the course. In Medicine, ratings A, B and C stood for successful progress as far as the first part of the Third Professional Examination, A with two or more Distinctions, and B with at least four merit certificates. D stood for progress only as far as the Second Professional, and E failure to reach this latter stage.

(*b*) *Results.*—The correlations between the various predictive measures and performance in secondary school and university examinations are set out in detail in Table 1.

Most of the correlations in Table 1 are low but positive. Excellence of university performance can, therefore, be predicted with some success about the age of eleven, but not at all accurately. The smallness of correlations in Table 1 is probably due in part to unreliability in the examinations which afford the basis of the criterion of university performance. There is also, of course, attenuation of coefficients through homogeneity, the standard deviation of quotients in this group being just over 8, instead of the normal 15. Correction

TABLE 1

CORRELATIONS BETWEEN SELECTION TESTS AND PERFORMANCE IN LEAVING CERTIFICATE AND UNIVERSITY EXAMINATIONS (CITY GROUP).

Correlations between and—		Arts (N=47)				Science (N=34)			
		SLC	U1	U3	U4	SLC	U1	U3	U4
	IQ_1	36	37	33	32	04	11	22	19
	IQ_2	58	50	46	44	– 02	18	– 04	02
	AQ	34	31	24	32	24	30	07	24
	EQ	14	37	33	21	– 01	– 13	– 17	– 28
	TE	32	33	20	14	17	– 14	– 15	03
	SLC	—	62	30	28	—	42	22	33
	U1	—	—	—	53	—	—	—	65
		Medicine (N=21)				All (N=102)			
		SLC	U1	U3	U4	SLC	U1	U3	U4
	IQ_1	14	17	19	31	22	25	26	27
	IQ_2	28	20	16	18	33	34	22	24
	AQ	– 01	24	33	24	25	30	20	28
	EQ	05	03	22	07	10	16	14	05
	TE	– 24	00	14	13	18	12	07	11
	SLC	—	24	17	28	—	48	24	31
	U1	—	—	—	71	—	—	—	60

(Decimal points omitted)

IQ_1, IQ_2, AQ, EQ: Scores from Moray House tests at age 11 plus.

TE : Teachers' estimates of primary school attainment.

SLC : Number of passes in Scottish Leaving Certificate.

U1, U3, U4: University performance up to end of 1st, 3rd and 4th year.

for homogeneity would raise the correlations, but such correction could hardly be justified here. As the assessments at age eleven show relatively high intercorrelation, the predictive power of a weighted battery in multiple correlation is only slightly higher than that of the best single test.

The correlations are generally higher in Arts than in Science or Medicine. Leaving Certificate results predict first-year performance in Arts quite well, perhaps because many first-year Arts students are working on subjects closely related to those studied at school. The results are consistent with the general finding that the best prediction of final performance (fourth year) comes from the latest assessment, that of first-year university performance. The appearance of negative correlations in the Science group, and the poor results from the English test at age 11, are disturbing. But it should be remembered that the number of cases is small; differences between correlations are generally not significant and too much should not be read into slight differences in predictive accuracy on this limited evidence. The field of investigation is of necessity insufficiently wide to provide a sound basis for deductions of general application.

Though these results support the general validity of the selection examination at age 11, certain cases at the extremes illustrate its fallibility in individual cases. There were ten students in this group whose I.Q. averaged from the two tests at age 11 on the Aberdeen standardization, was *below* 115. The university performance of nine of these ten students was much as one might expect. Four scraped through pass degrees in Arts ; one got a pass degree in Science ; one (I.Q. 114) graduated with Honours in Music ; one medical student (I.Q. 113) scraped through his Second Professional ; and two other medical students (both I.Q. 102) failed to do so. There remains, however, the tenth : I.Q. 112, Leaving Certificate performance mediocre, he passed both First and Second Professional Examinations in Medicine with Distinctions. At the other extreme, twenty-one students had average I.Qs. *above* 130. Eleven took Honours and seven (all women except one) Ordinary degrees in Arts or Science. One changed from Science to Arts and took an Ordinary degree. The remaining two are in Medicine : one passed both First and Second Professionals with Distinctions, while the other failed his Second Professional. Thus, among both high and low scorers in the tests at age 11, there are examples to warn against over-confidence in prediction at too early an age.

III.—ABERDEEN COUNTY FOLLOW-UP.

(*a*) *Selection system and tests.*—The organization of secondary schooling in the County of Aberdeen differs from that in the City. Though comparisons between Scottish and English systems are often misleading, it may be said that while the City, with its two types of secondary schools, providing five-year and three-year secondary courses, is nearer to the English system of secondary grammar and modern schools, the County follows to some extent a 'comprehensive' pattern.

In the County, allocation at the end of primary schooling is made, not to a type of school, but to a course of study—two language ; one-language ; commercial, technical, domestic (all with or without a foreign language) ; rural ; or modified. Each secondary school provides, if not all, at least a wide variety of these courses. This is designed to suit the geographical distribution of schools and pupils. Six schools in the County cater for pupils who stay on to the Leaving Certificate (taken usually towards the end of the fifth year of a secondary course). After three years there is centralization into the six schools (or, in special cases, to schools in adjacent counties) of pupils who wish to continue their course beyond the minimum school leaving age.

Allocation to courses at the end of primary schooling is made by a Promotion Board, consisting of the Director of Education, four members of the Education Committee and six teachers. The recommendations of the Board are based on scores in Intelligence, English and Arithmetic standardized tests, essay marks, teachers' estimates in English and Arithmetic based on work done in the primary school, head-teachers' summary assessments of the all-round ability of the pupil, and parents' wishes. Considerable weight is given to the last two. At no stage are any statistical procedures used : there is no summation of scores, no attempt to give a composite grading, no dividing line between admits and rejects and no hint of a border-zone. Choice of school is free, subject to the parents' willingness to meet extra transport costs and to the availability of the appropriate course in the chosen school.

Standardized tests were first introduced in the County for the 1950 promotion group. The pupils dealt with here were tested on one of three occasions : those promoted in January, 1950, were given tests MHT 38, MHE 15 and MHA 15 ; those promoted in April, 1950, were given tests MHT 39, MHE 16, and MHA 16 ; and those promoted in August, 1950, were given tests MHT 40, MHE 17, and MHA 17. The quotients from the parallel series of testings will here be treated as equivalent. That this procedure has its weakness is acknowledged, and it is justified only on the grounds that in the selection procedure, the equivalence of quotients from parallel tests is assumed.

(*b*) *Results.*—Of 2,388 County pupils who entered secondary schools in 1950, 866 entered courses recognized as leading to the Leaving Certificate. Of these 220 eventually entered the fourth year of the courses. In 1955 or 1956 (or on both occasions), 153 of these pupils were presented for the Leaving Certificate examinations. By the end of the session 1955-56, 77 pupils had gained passes which met the requirements for university entrance in Scotland.

TABLE 2

LEAVING CERTIFICATE PERFORMANCE AND TEST SCORES AT AGE 11 (COUNTY GROUP).

Performance in Leaving Certificate	Total Number	Mean IQ	Range of IQ	No. of IQ 120 and above	No. below IQ 110
Presented for examination	153	113	93—135+	36	60
5½ or more Higher passes	35	118	101—135+	14	5
5 or more Higher passes	51	116	98—135+	16	9
University entrance	77	116	98—135+	27	19
3 or less Higher passes	44	110	99—130	5	22

Correlation between Leaving Certificate performance (first attempt) and :	IQ	EQ	AQ
N=153	+0·36	+0·45	+0·27

Table 2 presents results in Leaving Certificate examinations in relation to performance in the intelligence test at age eleven. In assessing Leaving Certificate performance, the system was used (see IIa above) of counting two passes on the Lower standard as equivalent to one Higher pass. It is clear that there is a positive association between intelligence test scores at eleven and Leaving Certificate results five or six years later. Only five of the sixty pupils whose I.Q. at eleven was below 110 reached the high standard of 5½ Higher passes, and only five of the thirty-six pupils who scored above I.Q. 120 obtained three Highers or less. It is, however, also clear that a high I.Q. is no guarantee of success, any more than a low I.Q. is a harbinger of inescapable failure. This is shown by the very large overlap in range of I.Q. at different levels of Leaving Certificate performance. For university entrance, a minimum of four Higher passes (or equivalent) in certain specified groups of subjects is required. The majority of those who had scored above I.Q. 120 reached this standard, but so also did nineteen who had scored below I.Q. 110. Similar findings could be presented from any of the three transfer tests. The relations between Leaving Certificate results and the three tests are given as correlations at the end of Table 2.

Two cautions should be observed in viewing these results. Firstly, the year 1950 in which these pupils sat the transfer tests was the first year in which standardized tests had been used in the County. The intelligence test was the first applied, the English test second and the Arithmetic test last. Quotients were taken from the test manuals and were not re-standardized on the County population. Performance in the tests, particularly in the intelligence test, may have been affected by the absence of test-sophistication and of coaching among the pupils. Secondly, objection may be taken to counting the total number of passes in the Leaving Certificate, since passes gained in consecutive years may be added together. Weaker candidates in particular tend to divide their efforts over two years. However, if a stricter criterion is taken, of number of passes in the Leaving Certificate at the first attempt only, the results are similar. Among fifty-nine who obtained the equivalent of at least four Higher passes at the first attempt, the range of I.Q. (at age eleven) was 98 to 135+ ; twenty-one had scored above I.Q. 120; fourteen below I.Q. 110. Among the ninety-four pupils who obtained less than four Higher passes, the range of I.Q. was 93 to 130 ; fifteen had scored above I.Q. 120 ; forty-six below I.Q. 110.

Of the seventy-seven pupils who obtained university entrance qualifications, forty-nine entered Scottish universities, the University of Aberdeen receiving forty-six of them. It is of interest to see whether the promising performance of some of the original low scorers is being maintained at university level. The results are set out in Table 3.

TABLE 3

University Performance (First Year) and Test Scores at Age 11.

	Total No.	Range of IQ	No. of IQ 120 and above	No. below IQ 110
Entered Aberdeen University	46	101—135+	17	10
Passed all subjects	33	101—135+	12	6
Failed two or more subjects	8	102—131	2	4
Correlations between first-year university performance and—	IQ	EQ	AQ	SLC (1st attempt)
N=46	+0·18	+0·02	+0·17	+0·32

Of the nineteen who had scored below I.Q. 110 but gained university entrance, ten entered Aberdeen University. Six passed all their first-year degree examinations ; two failed in two subjects and two failed in three subjects The results in general show that, while the university performance of the original high scorers is above average and that of the original low scorers is below average, there can be no accurate prediction in any individual case. The correlations given in Table 3 are tetrachoric, as the data available at the time of inquiry were limited to passes and fails in the degree examinations. The number of cases is too small to permit any one of the tests at age eleven to be picked out as more, or less, accurate in prediction than any other.

IV.—Conclusions.

It is not the purpose of this report (nor, indeed, would it be diplomatic for the writers) to make comparisons between the contrasting systems of secondary education in the City and in the County of Aberdeen. Each system is designed to meet the special needs of its own area ; and any evaluation would need to take into account considerations far beyond the scope of this paper. These two differing systems provide complementary evidence on the relation of scores in transfer tests at age eleven plus to performance in secondary school and university examinations. The City data are better suited to statistical analysis: the County data afford evidence on pupils whose performance in transfer tests falls below what is usually taken as the border-zone. The field of inquiry is narrow and conclusions must be interpreted within the limits of the special groups studied.

The two follow-up investigations agree in showing that the transfer tests have a general validity in predicting scholastic attainment, but at best their accuracy of prediction can only be described as moderate. Though this report has chosen the intelligence test results for special attention, there is no consistent evidence to suggest that the intelligence test is less accurate or more accurate than the English or Arithmetic test, or that teachers' estimates would be any better. Assessment at age eleven could hardly be expected to achieve more than this modest level of prediction; for our criterion of success—namely, examination performance—is itself far from perfectly reliable, and many factors other than those measured affect academic achievement. Precise accuracy of prediction is, of course, not claimed for selection tests: rather is their prediction actuarial, the lower the score the less chance there is of success. In these terms the general validity of the tests is supported by the results. Yet frequent exceptions to the rule at both high and low levels of score emphasize the fallibility of early assessments and the need for a highly flexible organization of secondary school courses.

V.—ADDENDUM.

These results have been subsequently checked on parallel groups, with the additional assistance of Miss Kathleen Craib, M.A., Ed.B., Bristol Child Guidance Clinic, and James A. D. Michie, M.A., Ed.B., Aberdeen Grammar School. In Table 4 below, results are given for 1952 and 1954 university entrants from Aberdeen City, comparing eleven-plus performance with final (or fourth year) university performance. Table 5 summarizes results from the follow-up of County pupils who entered secondary school in 1951.

The negative correlation between English score at age 11 and performance in university science, reported above among 1953 entrants, was not found in the 1952 group, though it appears in the 1954 group. In the follow-up of the 1951 County group, the predictive value of the I.Q. is somewhat higher than in the 1950 group. Other results in general confirm the earlier findings.

TABLE 4

CORRELATIONS BETWEEN SELECTION TESTS AND PERFORMANCE IN UNIVERSITY FOURTH-YEAR EXAMINATIONS (CITY GROUPS).

Correlations between university performance and—	1952 university entrants				1954 university entrants			
	Arts	Science	Medicine	All	Arts	Science	Medicine	All
IQ$_1$	17	28	32	24	24	08	05	21
IQ$_2$	07	27	13	17	10	19	15	20
AQ	−04	30	24	13	−02	30	24	16
EQ	25	37	41	33	11	−16	13	13
TE	not available				22	15	21	23
SLC	52	39	14	38	38	14	11	29
(No. of cases)	(46)	(31)	(22)	(99)	(58)	(27)	(22)	(107)

TABLE 5

LEAVING CERTIFICATE PERFORMANCE AND TEST SCORES AT AGE 11 (COUNTY GROUP—1951 SECONDARY SCHOOL ENTRANTS).

Performance in Leaving Certificate	Total number	Mean IQ	Range of IQ	No. of IQ 120 and above	No. below IQ 110
Presented for examination	141	114	96—135+	41	57
5½ or more Higher passes	20	124	106—135+	15	2
5 or more Higher passes	35	121	103—135+	21	6
University entrance	57	118	98—135+	29	14
3 or less Higher passes..	43	108	96—135+	5	24
Correlation between Leaving Certificate performance (1st attempt) and—	IQ	EQ	AQ		
N=141	+0·53	+0·44	+0·34		

ACKNOWLEDGEMENTS.—These investigations would not have been possible without the generous facilities provided by the Education Authorities of the City of Aberdeen and of the County of Aberdeen, and by the University of Aberdeen, in permitting access to records and use of reports. In particular, acknowledgment is due to Mr. Clark and Mr. Johnston (City), Mr. Young and Mr. Gordon (County), Mr. George and Mr. Greig (University), and to teachers and head teachers.

TEST ANXIETY AND THE " ELEVEN-PLUS " EXAMINATIONS

By IRVING SARNOFF, SEYMOUR B. SARASON, FREDERICK F. LIGHTHALL, and KENNETH S. DAVIDSON

(*Department of Psychology, Yale University.*)

SUMMARY. After a brief review of our previous work in the area under investigation, the results of two studies which bear upon the validity of the Test Anxiety Scale (TA) are reported. Both studies involved the 11+ examinations. In the first study, it was predicted that TA scores would increase as children moved closer in time to the 11+ examinations. In the second study it was predicted that there would be a small but statistically significant negative correlation between TA scores and performance on the 11+ examinations. The obtained results failed to confirm either of these predictions. Drawing upon other empirical data which have been gathered as part of a research programme on the correlates of anxiety, several alternative interpretations of the findings are advanced.

I.—INTRODUCTION.

IN this paper, we shall present two studies which bear upon the validity of a questionnaire which was designed to measure the degree of anxiety which children feel in respect to school examinations. The first of these consists of an attempt to determine whether or not the imminence of the 11+ examinations is accompanied by a rise in the level of anxiety. The second study concerns the relationship between anxiety and performance on the 11+ examinations. In discussing the results, we shall raise a number of theoretical and methodological considerations which pertain not only to the data at hand but also the broader issues which they suggest.

II.—A BRIEF DIGEST OF PREVIOUS WORK WITH OUR MEASURES OF ANXIETY.

Before presenting the studies, however, we should like briefly to describe the particular scales which were employed as well as the major findings which have been obtained with their use in previous investigations. Such a preliminary review aims to provide a context within which the reader may be better able to evaluate the results of our present work.

The present studies are part of an ongoing project which was initially concerned with the development of a Test Anxiety Scale* (Sarason, *et al.*, 1957a). This scale, which now consists of thirty items, is group administered, and worded in the form of questions which each child answers by encircling either a ' yes ' or ' no ' alternative on an answer sheet. With children of school age, the scale is administered by an examiner who reads each question aloud to the group. The children then circle, on a separate answer sheet, the alternative which they choose. When the examiner is certain that all the children have answered a particular question, he proceeds to the next one. The questions in the TA scale involve attitudes toward and experiences in school situations. Three samples of these questions are :

Question 1—Do you worry when the teacher says that she is going to ask you questions to find out how much you know ?

---

* Copies of both the TA and GA scales for research purposes may be obtained free of charge upon written request to Professor Seymour B. Sarason, Department of Psychology. Yale University, New Haven, Connecticut, U.S.A.

Question 5—Do you sometimes dream at night that you are in school and cannot answer the teacher's questions ?

Question 24—When you are taking a test, does the hand you write with shake a little ?

The following findings with the Test Anxiety Scale (TA) are based primarily upon independent studies carried out in America over the past four years :

(1) The split-half coefficient of reliability is ·81 for a sample of 100 randomly selected school children in grades one, two and three.

(2) For children in elementary school (grades one through five), there is a significant and negative correlation between TA and IQ score (as measured by the Otis Alpha and Beta). The average Pearsonian correlation is approximately – ·23. The size and direction of this average correlation is very similar to that which was obtained with a sample of university students in a study (Sarason and Mandler, 1952), which related the Henmon-Nelson Test of Mental Ability to a TA scale similar in content but different in format from the children's scale described above.

(3) With school children there is a tendency for the correlation between TA and IQ to increase as grade increases.

(4) There is a significant and positive correlation between teacher's ratings of the child's anxiety and the child's own ratings of his anxiety, i.e., his TA score (Sarason, *et al*, 1957a).

(5) The more game-like the atmosphere in which an examination is administered (e.g., the Davis-Eells games) the less apparent are the interfering effects of test anxiety (Zweibelson, 1956).

(6) When he has experienced failure, the child who has a high TA score assigns blame to himself and others significantly more often than does the child who is low in test anxiety (Doris, 1957). Moreover, the child who is high in TA shows a stronger tendency to blame himself than to blame others, a finding compatible with that obtained in a similar study with college students (Doris and Sarason, 1955).

At a later stage of the project, a thirty-four item scale of general anxiety (GA) was developed with the same format and instructions as the TA scale. Examples of questions from the GA scale, which concerns anxiety and worry in a variety of situations, follow :

Question 7—Do you get scared when you have to walk home alone at night ?

Question 24—Do you worry that you might get hurt in some accident ?

Question 28—Are you afraid of being bitten or hurt by a dog ?

Like the TA, the GA appears to be a quite reliable measure, its split-half coefficient of reliability being ·81 for 100 randomly selected school children in grades one, two, and three. For large samples of school children the Pearsonian correlations between the TA and GA scales have tended to be of the order of +55. In a study of university students (Gordon and Sarason, 1955) using TA and GA scales adapted to their age level, we obtained a correlation between TA and GA which is similar in size and direction to that found with children.

In the past two years, the major focus of the project has been the study of thirty-two pairs of children matched for grade, sex, and I.Q., but differing in that one member of each pair was in the upper quartile on both the TA and GA scale (the high-anxious or HA subject), while the other member of the pair was in the first quartile of scores on both scales (the low-anxious or LA subject).

In each group, there was an equal number of boys and girls. The following are the principal findings which have thus far been obtained from these matched subjects :

(1) In an experimentally created learning situation, the performance of the HA subjects was significantly inferior to that of the LA subjects (Waite, *et al*, 1957). This result again coincides with one obtained on a similar study with college students (Mandler and Sarason, 1952).

(2) In a situation containing tasks which confronted subjects with relatively unfamiliar stimuli to which they were obliged to respond without the benefit of external support (i.e., the Rorschach cards) the performance of the HA group reflects the impact of interfering factors to a significantly greater degree that the performance of the LA group (Sarason, *et al*, 1957b).

The foregoing summary of the principal findings of our project would appear to support the validity of our scales*, especially in light of the similarity of findings in our studies of school children and university students. However, since the process of validating measures of so complex a psychological construct as anxiety must necessarily be a continuous one, we welcomed the opportunity to use the 11+ examinations as a way of observing the anxiety-evoking effects of a naturally occurring event which has been reputed to induce considerable anxiety amongst those who are required to contend with it. Moreover, we were interested in exploring the extent to which the correlates of our anxiety scales were similar in cultures which are quite different. That is, we wished to see whether or not such basic variables as grade and sex bore a relationship to anxiety which cut across certain cultural differences.

The results of our cross-cultural study are reported elsewhere (Sarnoff, *et al*, 1957). However, in order to lend further perspective to the findings about to be presented, it may be noted that the sex and grade correlates of test anxiety were of about the same order among English school children as among an equivalent group of American children. That is, in both countries, TA increases with age, and girls get higher scores on both the TA and GA scales than do boys. On the other hand, in view of the fact that the 11+ examinations have no counterpart in America, we predicted that while English school children (grades one through four) would not differ from American children (of the same school grades) on GA, they would have significantly higher TA scores. This expectation was confirmed.

III.—EXAMINATION STRESS AND TEST ANXIETY.

It appears reasonable to assume that the level of anxiety which a child possesses in regard to test-like situations should rise when that child perceives himself to be confronted by an actual examination. If this assumption is correct, a valid measure of test anxiety should reflect the rise in anxiety which occurs when the child is obliged to move temporarily closer to a concrete and stressful examination.

To assess this hypothesis, we decided to study the effects on TA scores of the imminence of the 11+ examinations. Specifically, we predicted that the imminence of the 11+ examinations should induce an increase in the TA scores of children who are on the verge of taking them.

In order to gather the data relevant to this specific prediction, it was, of course, necessary to administer the TA scales twice to the same group of subjects.

---

* We have not included findings pertinent to a third scale whose correlates we have briefly studied. This is the so-called Lie Scale, which consists of 11 questions that are intermingled with those comprising the GA scale. We constructed the Lie Scale in an attempt to assess the internal validity of the subjects' responses to both the GA and TA scales. The items in the Lie Scale concern fears or anxieties which are so universal that anyone denying ever having had such feelings could reasonably be presumed to be lying, or at least to be defending himself against the awareness or expression of a psychological fact.

The first administration occurred in June, 1956, with fourth graders of two English primary schools which are located in the Borough of Hendon, a residential district in the North of London. One of these schools caters largely to a lower-middle and skilled working class population. In the other school, about 50 per cent. of the pupils are from middle or upper-middle class families. Since June is one of the most carefree months in so far as formal academic requirements are concerned, it was felt that those TA scores should be relatively unaffected by immediately stressful examination pressures.

The second administration of the TA scale took place in early February, 1957, a week before the original group, now fifth graders, were scheduled to take the 11+ examinations. At the time of the re-administration of the TA scale, it was expected that the children would be pre-occupied with the anticipation of the 11+ examinations and this pre-occupation would be reflected by an increase in the group's TA scores as compared to the previous Summer. As can be seen from Table 1, our expectation was not supported. On the contrary, the mean TA scores for both boys and girls *decreased* significantly from June, 1956, to February, 1957.

TABLE 1

A Comparison of 1956 and 1957 Mean TA Scores Based upon a Random Sample of Twenty-four Boys and Twenty-four Girls from both English Primary Schools.

	Mean TA Scores in the 4th Grade, 1956	Mean TA Scores in the 5th Grade, 1957	Standard Error of the Difference Between Means	t	p
Boys	9·67	7·33	·269	8·69	..·001
Girls	12·33	9·38	·228	12·93	..·001

This unexpected finding puzzled us considerably until we discovered the following empirical fact which was not in our possession when the first study was inaugurated: *regardless of country, sex, or age, the mean TA scores of all groups—English and American—tend to decline from test to re-test.** In short, there appears to be a very distinctive ' order effect ' which characterizes re-test performance on the TA scale; and this effect seems to reduce the size of mean TA scores for re-tested groups.

This ' order effect ' will be discussed in detail in a subsequent article (Lighthall, *et al*, 1957). However, in anticipation we may say that re-testing could operate to put people on the defensive with respect to the emotionally loaded variable measured by the TA scale. Thus, upon retrospect, the re-tested individual may be more reluctant to acknowledge his anxious feelings. In this respect, we may be dealing with a phenomenon which psycho-therapists will immediately recognize: the fact that patients often appear to be more disturbed, and to express more disturbance, on their very first visit than on any subsequent interview.

---

* Lighthall, F. F., *et al* (1957) arrives at this conclusion as a result of using F-tests, following procedure in which each subject served as his own control. Also reported in their article are less sensitive statistical tests, employing a straight factorial design, which do not indicate a decrement in re-test TA means.

The ' order effect ' seems to have operated as effectively under the conditions of the English re-test as under the less stressful conditions of re-testing which have typified our work in American schools. Indeed, if our notions concerning the psychological forces which may produce the ' order effect ' are correct, we have reason to suppose that these forces should be more fully provoked by a re-test which occurs but a week prior to the 11+ examinations. Briefly, our psychological reasoning in this regard is as follows :

Since pupils are likely to be highly vigilant to all test-like situations a week before the 11+ examinations, they may well have perceived the TA questionnaires within the context of these examinations. In other words, the TA questionnaire may have been regarded by them as just as much a test as the 11+ examinations themselves.* If this were true, the pupils might have been inclined to be excessively cautious about admitting to behaviours which might, to a child's mind, be considered as ' bad ' by the school authorities. Thus, to answer ' Yes ' to items on the TA scale might have been tantamount to making a wrong or incorrect answer on a formal academic test.

One could say simply that the anxiety levels of the chidlren *actually did* decrease as they moved closer to the 11+ examinations. However, this interpretation sounds improbable in view of the clinical and anecdotal reports† concerning the disturbing effects of the 11+ examinations. Another possible interpretation might raise the question of the validity of the TA scale. It could hold, for example, that the 11+ examinations unquestionably do evoke anxiety. Hence, a valid measure of test anxiety would surely indicate a rise rather than a decrease in test anxiety. This interpretation would indeed, be difficult to deny if our findings lacked their definite trend. That is, if the test re-test comparison had merely indicated no difference in TA scores, we might be inclined to take this interpretation more seriously. However, the trend in the data is not only marked but also, as we have said, in line with other findings. On the other hand, if as we have postulated, a defensive reaction is set into motion as a consequence of simple re-testing, we are obliged to acknowledge that, under certain circumstances, the TA scale may be measuring defensiveness as well as test anxiety. One of the major reasons for interpreting the ' order effect ' as reflecting defensiveness stems from the consistent finding, to which we have already alluded, that boys get lower anxiety scores than girls. Elsewhere we have interpreted this finding as indication that, as a consequence of their cultural conditioning, it is harder for both English and American boys to admit to anxiety than for girls. In other words, we feel that greater defensiveness is aroused in boys in the process of answering the questionnaires than is aroused in girls.

---

* This line of reasoning tends to warrant serious consideration in view of the observations made by Miss Elizabeth Halsall, the person who administered the TA scales to both schools and on both occasions. Regarding the re-test, Miss Halsall notes that while some of the classes behaved in the same way as they did during the original testing, other classes appeared visibly more concerned with the procedure. These latter classes turned out to be the " B " and " C " classes of one of the schools ; i.e., children classified as having only average or below average intelligence. Commenting upon the change in the behaviour on the occasion of the re-administration of the TA scales, Miss Halsall writes . . . " They accepted the questionnaire uncritically, asked no questions about it, sought no elucidation of meaning or advice as to what to put, . . ., and questions that had raised a laugh last year were treated seriously, not to say earnestly, e.g., ' Do you get a funny feeling in your stomach ? ' . . . They've all got funny feelings in their stomachs ! Their general attitude was somewhat though not very tense."

† For example, in a letter to the senior author, the head teacher of one of the two schools involved in the present study noted that the pupils became so upset that they almost suffered ' nervous breakdowns ' while awaiting the results of their 1957 ' eleven-plus ' examinations.

IV.—TEST ANXIETY AND INTELLECTUAL FUNCTIONING.

Practicing clinicians and clinical researchers have long recognized the disabling effects which anxiety is likely to exert upon intellectual functioning. Moreover, these effects have been noted in a variety of intellectual tasks which range from the exercise of memory to the learning of new material, from the solution of simple problems to the creation of artistic works.

Generally speaking, the widespread intellectual impairment produced by anxiety appears to be a function of two of its psychological properties : (1) Its tendency to overshadow other feelings when it is aroused ; (2) The implacable demands which it makes upon the individual for its own diminution. In short, when the individual is made anxious, he seems obliged to deal, first and foremost, with his anxiety. Hence, he has difficulty attending to external stimuli, including whatever intellectual task may be at hand.

A certain peculiarity of both American and English cultures has made it difficult to conduct conclusive empirical studies of the relationship between anxiety and intellectual performances. We refer, of course, to the fact that an individual in these cultures is likely to interpret *any* intellectual task, indeed, any formally administered task whatsoever, as a test of *his worth as an individual*. Thus, it is virtually impossible to measure intellectual functioning without arousing, at the same time, anxiety which people possess with regard to being tested. To put this conclusion in technical terms, if the measure of any intellectual function elicits and reflects anxiety as well as the intellectual function in question, then the possibility of obtaining *independent* measures of anxiety, on the one hand, and intellectual functioning, on the other hand, is drastically reduced.

In our programme of research we have recognized and attempted to deal directly with this dilemma. Thus, we have been interested in ascertaining the extent to which anxiety, *which is specifically associated with test-like situations*, impairs intellectual performance in situations the test-like character of which is extremely clear as well as in situations which are much more ambiguous in so far as that character is concerned.

Since the 11+ examinations are endowed with an exceptionally high degree of test-like character, we decided to explore the relationship between performance on them and TA. Our expectation, based upon both theoretical considerations and the results of studies carried out in America, was that we would find a low and negative relationship between TA scores and 11+ results. Specifically, we predicted that children with low TA scores would perform somewhat more effectively on the 11+ examinations than children who are high in TA. We reasoned that the freer a child is from test anxiety, the less anxiety will tests stir up in him and, hence, the less will his intellectual functioning be impaired by the disabling effects of anxiety.

To test this specific prediction, we correlated the TA scores of our subjects with their subsequent performance, about a week later, on the 11+ examinations. The actual correlation which we obtained with an N of 152 boys and girls from both schools was − ·068, a finding which does not begin to approach statistical significance. Although this result stands in contrast to the significant negative correlations between TA and I.Q. which we have obtained with American samples, it may not necessarily be best interpreted as evidence against a relationship between TA and IQ in general. For it may well be that the absence of a statistically significant negative correlation reflects only the relationship between TA and this *particular* measure of I.Q. Indeed, it is conceivable that the extreme dread produced by 11+ examinations washes out whatever differential reactions high and low scores on the TA scale might reveal in other test situations which are, objectively, less important to them.

We had originally intended to select matched pairs of children in order to determine how individuals who were similar in 11+ score but markedly different in anxiety score would perform in other types ot problem-solving situations. If no significant differences were obtained, one would certzianly have grounds for questioning the validity of the TA scale. If significant differences did obtain, one would have a somewhat better idea of the role of task and instructional variables in relation to anxiety. For financial reasons it was not possible to follow through on the English studies. Fortunately, however, similar studies are being carried out in America on carefully matched pairs of subjects. The first two of these studies have been completed, and the rather clear-cut differences which were obtained offer support for the hypothesis that the disruptive effects of anxiety show up most clearly in situations where new learning is required (Sarason, *et al*, 1957b, Waite, *et al*, 1957). This hypothesis seems, furthermore, to be consistent with the psychological properties of anxiety to which we have alluded. For if anxiety makes inroads on the individual's ability to attend to external stimuli in an objective manner, its effect should be greater when the anxious individual attempts to deal with new and unfamiliar stimuli than with old and familiar ones. It may well be, therefore, that one of the major reasons the correlation between TA and I.Q. score is so low is that such scores are obtained from tests which contain familiar and over-learned content and which are accompanied by definite instructions not novel according to past experience. Such tests may arouse anxiety which does not have a strongly interfering effect because there are so many familiar aspects to the stimulus task which have the effect of quickly reducing anxiety. But we cannot overlook the individual who actually experiences much anxiety in a test-like situation—especially one like the 11+ examinations—but who, *in spite of strong anxiety* is able, somehow, to turn in an outstandingly good performance. One might quickly dispose of this phenomenon by saying that such an individual would probably perform even better if he were free ot anxiety. However, granting the cogency of this objection, it fails, nevertheless, to account for the fact that the individual has managed to deal simultaneously with both his anxiety and the task at hand.

But we have already postulated that anxiety makes very insistent demands upon the individual; that, when anxiety is strong, it diverts the individual's attention and energy away from external stimuli. How, then, is it possible for anyone ever to be extremely anxious *and* extremely effective on the very task which has aroused his anxiety?

Once again, in truthful answer to this rhetorical question, we are obliged to plead basic ignorance in so far as scientifically verified evidence is concerned. On the other hand, however, certain clinical observations do tend to point towards a meaningful hypothesis. This hypothesis holds that the highly anxious but intellectually effective individual has learned to compartmentalize his anxiety and, hence, to divert it into behavioural channels which run parallel to the ongoing task. Thus, it is possible that the anxiety is fed into a variety of somatic outlets while it is kept, at the same time, separate from the thought processes which must be applied to the task at hand.

Instances of this kind of ego defence may be found in respect to other kinds of tasks as well as intellectual ones. For example, many effective soldiers have been known to defecate involuntarily—so great is their panic during the heat of battle. Similarly, in running subjects through a variation of Asch's (1952) experiment on the yielding behaviour of persons who are placed in the position of a minority of one, Gerard (1956) reports that one of his subjects was smitten by a host of psychosomatic concomitants of anxiety. In fact, at one point, that subject asked to be temporarily excused from the task so that he could vomit.

Nevertheless, he did not once change his accurate judgment of the stimulus lines in order to conform to the inaccurate judgments of the minority.

The above discussion would suggest that research concerned with the adverse effects of anxiety should pay attention not only to variations in task and instructions but also to accompanying and delayed bodily reactions as well as to subsequent changes in perception of self and others. It is one thing to say that anxiety has had little or no effect on a particular problem-solving task ; it is another thing to imply either that one has adequately sampled problem-solving tasks or that other areas of functioning have likewise been little affected.

ACKNOWLEDGEMENTS.—The studies reported here were supported by a grant from the United States Public Health Service (M.712). Our data could not have been obtained without the technical aid and moral support of Professor P. E. Vernon of the Institute of Education, University of London, and Dr. N. F. H. Butcher, Borough Education Officer, Hendon. We wish to express our gratitude to these men and to the head teachers, teachers, and pupils of the schools studied ; people, who, owing to considerations of confidentiality, must remain anonymous. Thanks are also due to Mr. Paul Tacon, a post-graduate student at University College, London, who helped in the scoring and statistical analysis of the data. We are indebted to Miss Elizabeth Halsall, a post-graduate student at the Institute of Education, University of London, for her efficient administration of the scales and also for her help in translating our scales from 'American ' English into ' English ' English. Finally, we would like to thank Dr. Roger W. Russell, Executive Secretary of the American Psychological Association, who, in his former capacity as Chairman of the Psychology Department, University College of London, made office space and clerical facilities available to us throughout the Summer of 1956.

VI.—REFERENCES.

ASCH, S. E. (1952). Effects of group pressure upon the modification and distortion of judgments. In. G. E. Swanson, T. M. Newcomb, and E. L. Hartley (Eds.). *Readings in Social Psychology* (revised ed.). New York : Holt.

DORIS, J. (1957). Test anxiety and blame assignment in grade school children. Submitted : *J. Abnorm. Soc. Psychol.*

DORIS, J., and SARASLN, S. (1955). Test anxiety and blame assignment in a failure situation. *J. Abnorm. Soc. Psychol.*, 50, 335-338.

GERARD, H. B. (1956). Personal communication.

GORDON, E. M., and SARASON, S. B. (1955). The relationship between " test anxiety " and " other anxieties." *J. Personality*, 23, 317-323.

LIGHTHALL, F. F., DAVIDSON, K. S., WAITE, R. R., SARASON, S. B., and SARNOFF, I. (1957). The effects of serial position and time interval on two anxiety questionnaires. Submitted : *Educ. Psychol. Measmt.*

MANDLER, G. and SARASON, S. B., (1952). A study of anxiety and learning. *J. Abnorm. Soc. Psychol.*, 47, 166-173.

SARASON, S. B., and MANDLER, G. (1952). Some correlates of test anxiety. *J. Abnorm. Soc. Psychol.*, 47, 810-817.

SARASON, S. B., DAVIDSON, K. S., LIGHTHALL, F. F., and WAITE, R. R. (1957a). A test anxiety scale for children. In press : *Child Development.*

SARASON, S. B., DAVIDSON, K. S., LIGHTHALL, F. F., and WAITE, R. R. (1957b). Rorschach behaviour and performance of high and low anxious children. Submitted : *Child* Development.

SARNOFF, I., LIGHTHALL, F. F., WAITE, R. R., DAVIDSON, K. S., and SARASON, S. B. (1957). A cross-cultural study of anxiety amongst American and English school children. Submitted : *J. Educ. Psychol.*

WAITE, R. R., SARASON, S. B., LIGHTHALL, F. F., and DAVIDSON, K. S. (1957). A study of anxiety and learning in children. In press : *J. Abnorm. Soc. Psychol.*

ZWEIBELSON, I. (1956). Test anxiety and intelligence test performance. *J. Consulting Psychol.*, 20, 479-481.

BILINGUALISM AND NON-VERBAL INTELLIGENCE: A FURTHER STUDY OF TEST RESULTS

BY D. G. LEWIS

(*Department of Education, Queen's University, Belfast; Formerly of the Department of Education, University College of Wales, Aberystwyth.*)

SUMMARY. Previous investigations of the effect of (English-Welsh) bilingualism on performance in non-verbal tests of intelligence have not yielded a consistent pattern of results. In the present study a non-verbal test was given to ten-year-old pupils from primary schools in Wales. Linguistic background was assessed by means of a language questionnaire, the assessment being compared with teachers' ratings. Care was taken to select a homogeneous group of a thoroughly Welsh background (i.e., a bilingual group) together with a corresponding English-speaking group (with a background of no appreciable 'Welshness'). The test results showed a statistically significant difference, in favour of the monoglots, corresponding to about 8 points of I.Q. It is suggested that this difference may be due, in part, to (i) the test being a timed one, and (ii) the groups not being equal with respect to urban-rural differences. Again the 'verbal residue' of a non-verbal test might well be a non-negligible factor in measuring the intelligence of the bilingual child.

I.—INTRODUCTION.

PIONEER studies in mental development between Welsh-speaking bilingual children and monoglot English-speaking children were made by D. J. Saer (1922, 1923) and F. Smith (1923). Saer's investigations suggested that bilingual children from rural areas were at some disadvantage in their performance on the tests used, while Smith's results indicated a similar tendency with respect to children from urban districts also. Smith's tests were all verbal in character and were administered in English. Saer's tests were also largely verbal—mostly adapted from the Binet-Simon scale—and were administered in Welsh to the bilingual children. E. M. Barke (1933), however, used both a verbal and a non-verbal test and found a tendency for the bilingual children to be less successful on the verbal test only. A later study by E. M. Barke and D. E. Parry-Williams (1938) also obtained this result.

An important defect in these early investigations is the absence of any quantitative measure of the linguistic background of the children tested.* They do, however, suggest the possible advantage of using non-verbal tests in investigations of this kind.† However, in a more recent investigation, W. R. Jones and W. A. C. Stewart (1951) using a non-verbal test, together with an objective assessment of linguistic background, found a significant difference in favour of the monoglots. Yet, in a subsequent study by the former (1953)—in which the same test was used—no statistically significant difference was found. Finally, E. R. Morgan (1957), using three non-verbal tests obtained statistically significant differences (again in favour of the monoglots‡) in two tests, the performance of the monoglots being slightly superior also on the third (Raven's 'Matrices').

---

* Thus, children were merely divided into a bilingual (English-Welsh) group and a monoglot (English) group, and possible linguistic differences within the groups were ignored.

† E. M. Barke and D. E. Parry-Williams, for example, reported that the bilingual children did not have so great a vocabulary in either language as that of the monoglot children (ibid., p. 73).

‡ Morgan's grouping was actually three-fold, the group corresponding to the monoglots being described as ' of a predominantly English linguisitc background.'

In the present study Jenkins' Non-verbal Scale of Mental Ability was used as a measure of intelligence. The test is diagrammatic in character, and consists of eighty five items arranged in five sub-tests separately timed. Emmett (1949) has shown that, as part of a battery of verbal and spatial tests, the test has a high loading with the general factor (g) and also a small loading with the verbal factor (v). An alternative version of the test, with all written instructions in Welsh, was available for all children who preferred to use it.[*] Sixteen primary schools in the counties of Cardiganshire, Carmarthenshire, Merionethshire, and Pembrokeshire, were selected to secure a wide range of linguistic background, and all the ten-year-old pupils in these schools were tested.

II.—ASSESSMENT OF LINGUISTIC BACKGROUND.

The linguistic background of the children tested was assessed by means of a language questionnaire constructed along lines similar to Hoffman's Bilingual Schedule (1934), which had previously been administered to school children in Caernarvonshire (Miles, 1953). The first half of the questionnaire consists of sixteen questions of the type 'Does your mother speak to you in Welsh?,' 'Do you speak to your father in Welsh?', 'Does your father speak to your mother in Welsh?,' etc. All answers have to be selected from the five alternatives: *Always*, *Mostly*, *About half the time*, *Not often*, and *Never*. Fourteen of the questions relate directly to the speaking of Welsh in the home, the remaining two being 'Do you speak Welsh to your friends?' and 'Do your friends speak Welsh to you?'

In the second half of the questionnare each of the questions is repeated with 'Welsh' replaced by 'English.' (Thus, for example, the question 'Does your mother speak to you in Welsh?' is repeated in the form 'Does your mother speak to you in English?') By checking that answers to alternative forms of the same question correspond, and disregarding any answers which do not, a high reliability for the questionnaire is ensured.[†]

The questionnaire is scored on the number of acceptable answers. At one extreme a child answering *Always* to all the questions answered[‡] (in the first half only) would obtain a percentage score of +100, while, at the other extreme, a child answering *Never* to all these questions would obtain a percentage score of −100. All children, therefore, from homes where Welsh is habitually spoken, would have scores of, or near, 100; while all from homes where Welsh is seldom, if ever, spoken would have scores of, or near, −100. Questionnaire scores not near either of these limits would represent intermediate grades of linguistic background.

---

[*] The oral instructions during the administrations of this version were also given in Welsh.

[†] The validity of the questionnaire, as a measure of 'Welshness' in linguistic background, is a more complex matter, and one that can hardly be settled statistically since there is no agreed criterion as to what constitutes a (quantitative) measure of 'Welshness.' Thus one could criticize the narrow content of the questionnaire. (Account is taken only of the actual speaking of Welsh, and no questions relating, for example, to the reading of Welsh books in the home, or to the listening to Welsh programmes on the wireless, are included.) Against this, however, it could be fairly maintained that the questionnaire does concentrate on the most important single factor of the child's linguistic background, namely the active use of Welsh in the home.

[‡] For children with no brothers or sisters, or with only one parent living the number of questions answered would be less than sixteen.

III.—SELECTION OF GROUPS.

In the report of the questionnaire as administered in Caernarvonshire, all who scored from $+51$ per cent. to $+100$ per cent. were grouped together as ' definitely Welsh '* A recent investigation of the comparative attainments of bilingual children, in which the same questionnaire was used, also employed this grouping (W. R. Jones, *et al.*, 1957). However, a detailed examination of the questionnaire answers obtained in the present study showed considerable differences in background within a group so defined.† Comparisons in test performance with other groups—and, in particular, with the corresponding English group—would be more informative if the Welsh group was homogeneous and of a thoroughly Welsh background. The Welsh group was, therefore, restricted to those who scored from $+90$ per cent. to $+100$ per cent. on the questionnaire. Similarly, a selective and homogeneous English group was ensured by restricting this to those who scored from -90 per cent. to -100 per cent. on the questionnaire. (It was found that both groups could be selected in this way without a serious depletion of numbers.) Finally, the remaining children were divided into two groups by a division at the questionnaire score O—the theoretical point of separation between a predominantly Welsh and a predominantly English background.

Agreement with Teachers' Ratings.—All the children tested were rated by their teachers on a four-point scale as follows :

W,	children whose first language is Welsh and whose knowledge of English is slight ;
We,	children whose first language is Welsh, but who can express themselves with a fair degree of facility in English ;
Ew,	children whose first language is English, but who can express themselves with a fair degree of facility in Welsh ;
and E,	children whose first language is English and whose knowledge of Welsh is confined to that learnt in school.

The lack of symmetry in the category descriptions W and E is due to the fact that there would be few, if any, ten-year-olds whose knowledge of English is confined to that learnt in school. While all in category E might be expected to have a completely English background (i.e., a questionnaire score of -100 per cent.) a corresponding statement about those in category W would not be true. Indeed, some Welsh-speaking children may well have been placed in this category irrespective of the degree of ' Welshness ' of their background, because of their limited command of English. From another point of view, however, the wording in the description of this category (children " whose knowledge of English is slight ") may have been too restrictive, i.e., children of a thoroughly Welsh background may well have been excluded merely because they are able to express themselves with reasonable fluency in English. The teachers' ratings are clearly dependent upon fluency in the second language no less than upon the language of the home, and exact agreement between them and the questionnaire grouping should not be expected. The correspondence obtained is shown in the following table :

---

* Ibid., pp. 20-21.

† Thus while most children in such a group come from homes where Welsh is always spoken by all in the family, many come from homes where this is not so. Only one parent, for instance, might be constantly speaking in Welsh to the child, or there might be a brother or sister with whom the child would speak in English *About half the time*. Such children, too, would often have friends with whom Welsh would only be spoken *About half the time* or even *Not Often*.

		Teachers' Ratings				
		W	We	Ew	E	Totals
Questionnaire Groups :	1. (+90% to 100%)	20	73*	—	—	93
	2. (0% to +89%)....	6	84	5	4	99
	3. (−1% to −89%)..	1	20	21	53	95
	4. (−90% to −100%)	—	3	3	82	88
	Totals	27	180	29	139	375

IV.—DESCRIPTION OF GROUPS.

The characteristic features of the groups, as revealed by an examination of the questionnaire data, are as follows :

Group 1.—The language in the home of all in this group is entirely Welsh. Not only is Welsh always spoken between each parent and the child, but also between the child and his brothers or sisters. Welsh is also spoken to friends either '*mostly*' or '*always*.'

Group 2.—Here the home conditions are more varied, but in no case is Welsh always spoken between all members of the family. [Thus, while one or both of the parents might speak to the child in Welsh ('*mostly*' or even '*always*'), English is often spoken between the child and his brothers and sisters. Sometimes, too, one of the parents cannot speak Welsh, and in almost all cases less Welsh is spoken to friends than is spoken in the home.]

Group 3.—In this group, too, the home conditions are varied. Most of the children from homes in which only one parent is Welsh-speaking fall in this group. Thus, while that parent does speak Welsh to the child (usually '*not often*' or '*about half the time*') no Welsh is ever spoken by the other parent.

Group 4.—Here the language of the home is English, and, in almost all cases, Welsh is never spoken by the child. Indeed, most children in this group have questionnaire scores at the limit of −100 per cent., i.e., a completely English background. For other children the slight departure from this limit is usually due to Welsh being spoken ('*not often*') by friends.

The above order of description is thus one of decreasing 'Welshness' of background. It is important to realize, however, that the extreme groups 1 and 4, although symmetrically placed in the range of questionnaire scores, do not correspond with respect to the actual linguistic ability of the children. Thus all in group 1 can speak English, whereas none in group 4 can speak Welsh. Group-4 children are all monoglot ; group-1 children are bilingual. The order of group description can thus be regarded as an order of decreasing bilingualism no less than one of decreasing 'Welshness' of background.

Differences in age between the groups were found to be negligible.† In each group the proportion of boys to girls was practically the same. A comparison of parental occupation revealed no differences between the groups in socio-economic status.

---

* These 73 children showed a much higher attainment on a test of English usage (Moray House English Test 21) than the remaining 20 children in group 1, their mean score being 27·99 as compared with a mean score of 14·20 for the other 20 children. It seems reasonable to suppose, therefore, that most of these children were not rated W merely because of their command of spoken English. It is also relevant to note that the mean score of the 7 children rated W who are *not* in group 1 was only 9·86,—a fact which strongly suggests that it is their limited fluency in English which accounts for them being rated W.

† The mean ages of the groups (to the nearest month) are 10 : 5 for group 1, 10 : 4 for group 2, 10 : 5 for group 3 and 10 : 4 for group 4.

V.—TEST RESULTS.

The mean scores of the groups, together with the standard deviation of scores within each group, on the Jenkins Non-verbal Scale of Mental Ability, are shown below:

Group	1	2	3	4
Mean	33·30	38·36	38·94	40·98
S.D.	16·37	15·95	16·26	17·69

The groups are thus not even approximately equal in test performance, the order of decreasing 'Welshness' of background being also one of increasing mean score. In particular the mean score of group 4 is considerably higher than that of group 1, the difference corresponding to about 8 points of I.Q., according to the norms obtained in the original standardization of the test by the National Foundation for Educational Research.* The over-all differences between the four group means are significant at the 5-per-cent. level (F=3·57 for 3 and 371 degrees of freedom); the difference between the means for groups 1 and 4 is significant at the 1-per-cent. level (t=3·02 for 179 degrees of freedom).

VI.—DISCUSSION.

The statistical significance of the group differences merits further consideration. As the groups were selected solely on the basis of linguistic background, it would seem reasonable, *prima facie*, to expect no bias in general mental ability—so that any obtained group differences would be attributable to the random effects of sampling. But statistical analysis has shown this explanation to be improbable. It is far more likely that the higher mean score of group 4, as compared with group 1, reflects a real difference between monoglot and bilingual children in performance on the test. In this respect, the results support those of Morgan's investigations and prompt consideration of why ten-year-old bilingual children should appear at a disadvantage on a non-verbal test of intelligence.

One reason—though possibly providing only a partial explanation—is apparent from Morgan's results in that of the tests he employed the one in which the retardation of the bilingual children was least (and non-significant) proved to be the one test without a time-limit, namely Raven's 'Matrices.' As has been suggested in a recent paper (Morrison, 1958), it may well be that the bilingual child, having to make a choice between two languages (and possibly trying to use both on occasions), tends to be slightly slower at thinking than the monoglot, and is thus penalized on all timed tests.

Performance on intelligence tests is, of course, influenced by the factor of socio-economic status and also that of urban-rural differences. As has been stated the groups of this investigation did not differ appreciably with regard to parental occupation (though of course a more thorough analysis of the socio-economic factor might have revealed some differences). With regard to the urban-rural factor the groups were not equal, a slightly higher proportion of the Welsh-speaking bilinguals coming from a rural environment than of the English-speaking monoglots. This may well have had some effect on the test results.

* The mean scores for groups 1 and 4 correspond to quotients of 97 and 105 respectively.

BUREAU OF EDN. & PSY. RESEARCH
27-8-70
LIBRARY
J 222

Finally, although the intelligence test used in the present investigation is non-verbal in type language is inevitably involved in the test administration (if not also in the thinking processes of the testees). The small but appreciable loading of the test with the ' v ' factor of Emmett's analysis may be noted in this connection. The influence of this ' verbal residue ' in non-verbal tests could well prove to be far from negligible in assessing the intelligence of the bilingual child.

VII.—References.

[1] Barke, E. M. (1933). A study of the comparative intelligence of children in certain bilingual and monoglot schools in South Wales. *Brit. J. Educ. Psychol.*, III, 237-250.

[2] Barke, E. M., and Parry-Williams, D. E. (1938). A further study of comparative intelligence of children in certain bilingual and monoglot schools in South Wales. *Brit. J. Educ. Psychol.*, VIII, 63-67.

[3] Emmett, W. G. (1949). Evidence of a space factor at 11 + and earlier. *Brit. J. Psychol.*, (*Stat. Section*), II, 3-16.

[4] Hoffman, M. N. H. (1934). *The Measurement of Bilingual Background.* New York: Columbia University.

[5] Jones, W. R. (1953). The influence of reading ability in English on the intelligence test scores of Welsh-speaking children. *Brit. J. Educ, Psychol.*, XXIII, 114-120.

[6] Jones, W. R., and Stewart, W. A. C. (1951). Bilingualism and verbal intelligence. *Brit. J. Psychol.* (*Stat. Section*), IV, 3-8.

[7] Jones, W. R., Morrison, J. R., Rogers, J., and Saer, H. (1957). *The Educational Attainment of Bilingual Children in Relation to their Intelligence and Linguistic Background.* Cardiff: University of Wales Press.

[8] Miles, T. R. (1953). *Bilingualism in Caernarvonshire.* Cardiff: University of Wales Press.

[9] Morgan, E. R. (1957). *Bilingualism and Non-verbal Intelligence: A Study of Test Results.* Pamphlet No. 4. Collegiate Faculty of Education, Aberystwyth.

[10] Morrison, J. R. (1958). Bilingualism: some psychological aspects. *The Advancement of Science*, 56, 287-290.

[11] Saer, D. J. (1922). An enquiry into the effect of bilingualism upon the intelligence of young children. *J. Exp. Ped.*, VI, 232-240, 266-274.
—— (1923). The effect of bilingualism on intelligence. *Brit. J. Psychol.*, XIV, 25-38.

[12] Smith, F. (1923). Bilingualism and mental development. *Brit. J. Psychol.*, XIII, 271-282.

PERSONALITY FACTORS IN THE FORMATION OF OCCUPATIONAL CHOICE

By SHEILA M. CHOWN

(*Department of Psychology, Liverpool University.*)*

SUMMARY. 1.—Ninety-six boys and ninety-six girls selected from the third to seventh forms of eight grammar schools were given a questionnaire about their interests, behavioural traits, and reasons for choosing their desired occupation. Intelligence test scores were available for half the group.

2.—Interests were compared where possible with the type of jobs desired; practical spare time interests were related to job choices which involved these interests.

3.—Intelligence was found to have a limiting effect on choice; the most intelligent in this group made a free choice, but those of low intelligence tended to choose office work.

4.—The answers to the personality questionnaire were factor analyzed, giving two unrelated factors of confidence and sociability. Choice of office work was related to low confidence scores in the case of girls, and high sociability scores in the case of boys.

5.—The reasons for choice given by the subjects were factor analyzed; two unrelated factors were isolated. Vocation was related to the choice of occupations requiring further training, and unadventurousness to choice of office work.

6.—The importance of these results in relation to choice of occupation is discussed.

I.—INTRODUCTION.

STUDIES of influences shaping occupational choice have usually been confined to those factors which are external to the person concerned. Few attempts have been made to link the choice of a particular type of occupation to the characteristics of the person himself. The present investigation sought for relationships between values, interests, ability, and vocational plans, among a sample of grammar school children.

II.—PREVIOUS RESEARCH.

Both American and English investigators (Fleege and Malone, 1946; Ginzberg, *et al.* 1951; Jahoda, 1952) have reported that, though ability and type of career chosen may become more compatible with increase in age, too little attention is paid to ability even among older children.

Spare time interests have more often been investigated among those already well established in a career than among those about to choose a job (Strong, 1943). Inventories given for vocational guidance purposes generally deal with occupational rather than spare time interests. However, Berdie (1943) reported that hobbies were related to expressed and measured vocational interests among 136 American freshmen.

Since the present work was carried out, Cattell (1956) has published a report showing that personality profiles based on his sixteen-factor personality questionnaire will distinguish between those already in widely differing occupations.

---

* Based on part of a thesis submitted for the degree of Ph.D. in the University of Liverpool. The writer wishes to express her gratitude to the staff of the department, particularly to Dr. Alastair Heron and Mr. Dermot Straker, and to the Heads of Schools and Youth Employment Officers who made the research possible.

Values have been studied by M. D. Vernon (1937 and 1938) among schoolgirls and university students. The motivation behind their choice of careers varied from altruism to desire to have a good time. She did not attempt to link the motives to the actual work chosen. Fleege and Malone (1946) attempted to classify reasons given for choice under various types of value heading, and Ginzberg maintained that values become important in choice at 15-16 years-old. All studies investigating altruism have found it to be a more frequent motive among girls than among boys.

III.—SUBJECTS.

Two education areas were sampled in the enquiry: they will be called Area X and Area Y. Area X has a population of about three-quarters of a million and is a port and commercial centre. It includes residential districts of many types. The area has fifteen maintained grammar schools. Area Y, with a population of approximately 100,000, serves as a dormitory for people who work elsewhere, and is predominantly middle-class. However, the area does contain docks and factories and has its 'working-class' district also. It has five maintained grammar schools.

Two boys' schools and two girls' schools were chosen in each area from the non-denominational ones available. Twenty-four children were selected from each school; the age range was thirteen to eighteen, and children were drawn from the third to the seventh years of the schools. More children were seen in the fifth form than in other forms, since this was the period when some children were about to leave and others to stay on. The number of children to be seen in each 'year' and each form within the 'year' was specified, and the subjects were then chosen according to their surname initial.

IV.—METHODS OF STUDY.

The subjects completed a forty-five minute questionnaire under the writer's supervision, and were later interviewed individually. Information was obtained about intelligence from tests taken at 11. The 103 subjects still at school a year after the first interview completed a follow-up questionnaire.

A.—Interests.—A paired comparison test of spare time interests was devised for the questionnaire, using five groups of interests. These groups were Aesthetic, Intellectual, Practical, Physical, and Social. Three widely different examples of each interest were used. Thus, the Physical group contained 'Play in a team game,' 'Go swimming,' and 'Attend a gym. display.' Subjects were asked to say which *group* of things they would rather do out of each possible pair. This test was included in the follow-up questionnaire. Subjects were also asked to say how they would spend an unexpected day's holiday, and it was found that their replies, classified into the same five groups, agreed with the first or second test choice in 80 per cent. of the cases.

B.—Intelligence.—Owing to the short time available for testing, it was thought better to rely on the 11+ Moray House intelligence test results, available for half the subjects, rather than to attempt to get a new estimate of intelligence.

C.—Personality.—No existing test could be found which was short enough and which would give the subject's own expressed opinion about some of his personality traits. At the beginning of the study the investigator was thinking in terms of Sociability, Sympathy, Persistence, Tolerance, Energy, Contentment, and Self-confidence. Before proceeding, possible relationships between these traits were sought in the list of personality trait clusters given by Cattell (1943). It was found that very little overlapping occurred between them, except in the case of ' energy ,' which appeared to overlap with ' self-confidence ' and ' persistence.' In view of this, the writer felt justified in attempting to use the seven traits as units in making up a personality questionnaire. For each of the traits a description of a person who lacked it and one of a person who possessed it in an extreme form were written. Each pair of descriptions was then divided into three contrasting sets, giving twenty-one pairs of sentences in all. There was no guarantee that the descriptions of each trait corresponded with the way in which the raters in Cattell's experiment had used the terms ; nor was there any guarantee that the three pairs of sentences actually described behaviour due to the same trait. However, for purposes of arrangement, they were not placed next to each other, nor next to any of the other traits more than once in the questionnaire. Each subject was asked to read through each pair of sentences and to tick the alternative which he felt most resembled his own behaviour. The personality questionnaire was also included in the follow-up.

D.—Reasons for choosing a job.—This section consisted of the twenty-four main reasons for choosing jobs given by subjects in previous research—particularly in the findings of Jahoda (1952), Valentine and Ritchie (1927) and Fleege (1946). Each ' reason ' was painted in black in one-inch high block capitals on a strip of cardboard, and held up in front of the groups of subjects who were in this way forced to make an immediate decision about it. Each subject was asked to put a tick under the ' Yes ' column of the answer sheet provided if the reason was one which had influenced him, and under the ' No ' column if it had not been an influence.

V.—RESULTS.

A.—Interests.

(i) *Interests and job preference.*—Many occupations were chosen by only a few people. In order to compare different jobs, it was decided to classify them according to the job interests involved in them. The Kuder (1939) classification was used and ' translated ' into the categories of the present interest test (See Table 1). Occupations listed under any heading(s) of the Kuder scheme could then be directly compared with each other, or examined for their correspondence with the subjects' spare time interests. A comparison of the interests involved in the job preferences and the spare time interests showed that the two by no means corresponded. Spare time interests were less ' intellectual ' and ' practical ' and more ' physical,' ' social ' and ' aesthetic ' than those involved in the jobs. However, the coincidences of spare time interests and job interests were counted, and where the expected frequencies were large enough, chi-squared tests were applied. The only significant relationships found were those between practical spare time interests as measured on the test, and jobs which involved practical interests ($P<\cdot05$ for boys, and nearly ·05 for girls). Girls with physical or athletic interests also seemed likely to choose ' practical ' jobs ($P<\cdot05$). Unfortunately, there were too few choices involving physical and aesthetic interests to see whether people desiring these jobs had the corresponding spare time interests.

(ii) *Change of interests with time.*—It was found that twenty-eight boys out of fifty-six, and twenty-six girls out of forty-seven, had changed the order of their interests in one year to the extent that the rank order correlation was

TABLE 1

RELATIONSHIP OF KUDER INTEREST CLASSIFICATION TO LEISURE INTERESTS.

Kuder Occupational Interests	Present Leisure Interests
Social Service Persuasive	Social
Literary	Intellectual
Computational Scientific Clerical	Intellectual and Practical
Mechanical	Practical
Art Music	Aesthetic
—	Physical

·7 or less. However, comparatively few major reversals of interests occurred—in only a quarter (15/56) of the boys, and a third (17/47) of the girls. Since interests are held to fluctuate most during the period of adolescence, the amount of change found after a gap of one year was less than might have been expected in this 13 to 18-year age group, and the 'interest test' appears to be valid and reliable as a guide to spare time interests.

B.—Intelligence.

(i) *Intelligence and job choice.*—The scores of most of the subjects in one geographical area on the Moray House intelligence tests used in the 11+ selection were available. The slight differences between the average scores at different levels in the schools were found not to be significant.

The sample was divided into thirds approximately, on the basis of their 11+ scores, and the number choosing a career involving a university training was counted for each third. The result of a chi-squared test between desire for such training and the intelligence test results was significant ($P < \cdot 02$), but the greatest contributions to chi-squared came from the third with lowest scores, where fewer people chose university training than would be expected by chance.

The job choices were also divided into office work (or the equivalent) and 'other work.' Again there was a significant difference in the level of intelligence of these two groups ($P < \cdot 05$). The greatest contribution to chi-squared came from the proportion of the low-intelligence group choosing office work, which was larger than expected by chance. It seems that 'intelligence' acted in a negative way on choice of career; those above a certain level chose freely, but those below it tended to choose office work.

C.—Personality and reasons for choice.

The answers to the personality questionnaire and the suggested reasons for choice were factor analysed. Unweighted factor scores were then computed for each subject. As a result, it has been possible to look for relationships between the factors themselves and between the factors and job choices.

(i) *Personality.*—Three questions were answered negatively by less than 25 per cent. of the subjects and were, therefore, omitted from the factor analysis. There were certain sex differences in the numbers answering some of the questions positively, and so it was decided to include sex as a variable in the analysis. The effect of age on the answers to the questionnaire was also examined but no significant correlations were found. The tetrachoric inter-correlations of the answers to each question were calculated. The matrix was factor analysed by Thurstone's centroid method and three factors were taken out, which together accounted for 30·6 per cent. of the variance. Orthogonal axes were used and these were rotated to obtain simple structure.

Two factors appeared to be meaningful. Factor I consisted of the positive answers to seven questions. It has been tentatively called 'confidence.' The loadings of these variables on Factor II were all under ·16.* Male sex had a loading of ·41 on Factor I.

TABLE 2

PERSONALITY.

FACTOR I

Number	Question	Loading
20	" Rarely wants to change his mind "	·78
5	" Full of energy and enthusiasm "	·59
7	" Goes on working in the same way whatever difficulties occur"	·52
19	" Likes to go on with a thing till finished and can't bear interruptions "	·44
13	" Sets out to enjoy everything he does "	·43
12	" Never has any doubts about it after making a choice "	·37
14	" Will not give up a course of action once he has decided on it "	·32

Factor II was made up of six variables, and has been called 'sociability.' The loadings of all its variables on Factors I and III are under ·15. Female sex had a loading on Factor II of ·26.

TABLE 3

FACTOR II

Number	Question	Loading
16	" Soon becomes the centre of a group and enjoys social life "	·70
11	" Gets to know other people easily."	·66
1	" Talkative and cheerful "	·54
9	" Hurt when other people are hurt "	·48
17	" Enthusiastic and full of plans for the future "	·41
15	" Finds it extremely distressing to see others unhappy "	·36

Factor III appeared to be a mixture of 'confidence' and 'discontentment' items, and was not investigated further.

---

* On Factor III, numbers twelve and fourteen had loadings of ·39 and ·45 and these would have been omitted when scoring Factor I had Factor III been used.

Each subject was assigned two 'factor scores,' according to the number of relevant questions positively answered. These two scores were found to be uncorrelated, r being – ·02. The two factors fit in fairly well with the original classification. 'Confidence' contains most of the original 'persistence,' 'energy,' and 'self-confidence' statements. 'Sociability' is made up of the original 'sociability' statements and two of the 'sympathy' ones. Only the sentences originally labelled 'tolerance' and 'contentment' have been of no assistance in building up the factors.

(ii) *Change in personality scores after one year.*—The personality questionnaire was completed again by 103 of the original subjects, after an interval of one year. The responses were scored for the two factors. A comparison of the factor scores on the two occasions gave a product-moment correlation of ·39 for 'confidence' and of ·72 for 'sociability.' The consistency of 'confidence' was found to increase linearly with age but did not become reasonably consistent until the highest age group. 'Sociability' was answered more consistently at all ages, but the consistency fluctuated in an irregular manner. There was no evidence of curvilinearity, and the consistency of boys and girls did not differ significantly at any age tested.

(iii) *Reasons for choice.*—Three reasons were rejected by over 90 per cent· of the subjects and were omitted. These were 'I had to leave school now so had to decide on something,' 'It is only a stop-gap,' and 'I took the job to be sure of having one.' There were again sex differences in the answers to the questions, so sex was added to the correlation matrix as an additional variable. The relationship between age and reasons accepted was found to be significant in two instances—'It's a secure job,' and 'There is opportunity for promotion.' Age was partialled out from the intercorrelations of these variables. Tetrachoric correlations were used in the matrix, which was analysed by the centroid method. The three factors obtained accounted for 34 per cent. of the variance. The orthogonal axes were rotated to obtain simple structure. Factor I contained five reasons which all had loadings of less than·25 on Factors II and III Female sex had a loading of ·44 on this factor, which has been called 'Vocation.'

TABLE 4

REASONS FOR CHOICE.

FACTOR I

Number	Reason	Loading
5	" I want my life to be useful "	·71
12	" I want to do something to help other people "	·67
8	" I have always been extremely interested "	·65
24	" I feel I was 'meant' to do work like this "	·60
10	" I want to meet new and exciting people "	·44

Factor II contained seven of the reasons, the loadings of which were all less than ·25 on Factor I and, except for number 14, less than ·27 on Factor III. Sex did not appear to be associated with Factor II. This factor seems to represent the desire for security and safety and has been named 'Unadventurousness.'

TABLE 5

FACTOR II

Number	Reason	Loading
3	" The money is good "	·63
7	" It is a secure job "	·60
18	" I want a job to help my family "	·55
16	" Jobs like this were easy to get "	·50
14	" It is near home "	·48
1	" My parents want me to do work like this "	·48
2	" The hours are reasonable "	·48

Factor III consisted of seven items all related to ease of getting the job ; three of the items were negative, i.e., not ticked, and partly for this reason, the factor was not used. Scores on I and II were uncorrelated, r being − ·02.

(iv) *The relationship between ' personality ' and ' reasons for choice.'*—The scores for all the subjects on each of the ' personality ' factors were correlated with their scores on each of the ' reasons ' factors. The differences between the coefficients for boys and girls were not significant. However, the relationship between factors varied according to sex. ' Confidence ' appeared to be positively related to ' vocation ' in the case of girls only (r= ·25), and negatively to ' unadventurousness ' in the case of the boys only (r= − ·31).

(v) *Variability of factor scores with age.*—Means for the boys and girls of various ages on each of the factors are shown in Figure 1. Though these means alter with age, they do not always do so in the manner expected by common sense.

At all ages boys have higher mean ' confidence ' scores than girls. The scores tend to fall with age for boys ; for girls the scores fall until sixteen and then increase. On the ' sociability ' factor, girls at all ages score more highly than boys. The means alter only slightly with age ; they tend to increase until fifteen (for boys) and sixteen (for girls) and then to fall. Boys' and girls' mean scores for ' unadventurousness ' closely resemble each other. The scores increase until fifteen or sixteen, then fall off fairly rapidly. It is hardly surprising that the trend away from ' safe ' reasons should rise as subjects increase in age, and in maturity. Another cause of the decrease may be that those making for a ' safe' job leave at sixteen ; on the other hand, those who remain often desire jobs (such as teaching) which other people tend to regard as unadventurous and safe. Apparently it is not ' unadventurous ' reasons which prompt these subjects to aim for them. At all ages girls score more highly then boys on the ' vocation ' factor. For both boys and girls, the pattern of scores rises at fifteen, drops at sixteen, and then rises. Since sixteen is an age at which many subjects must make a final choice of job, it seems likely that this fact may be related to the drop.

The variances in different age groups on each factor were also studied but showed no meaningful significant trends.

(vi) *Relationship between factor scores and present choice of job.*—Owing to the comparatively small numbers who chose exactly similar jobs, fairly wide categories for job-choices had to be used. These were for girls, ' Office or office

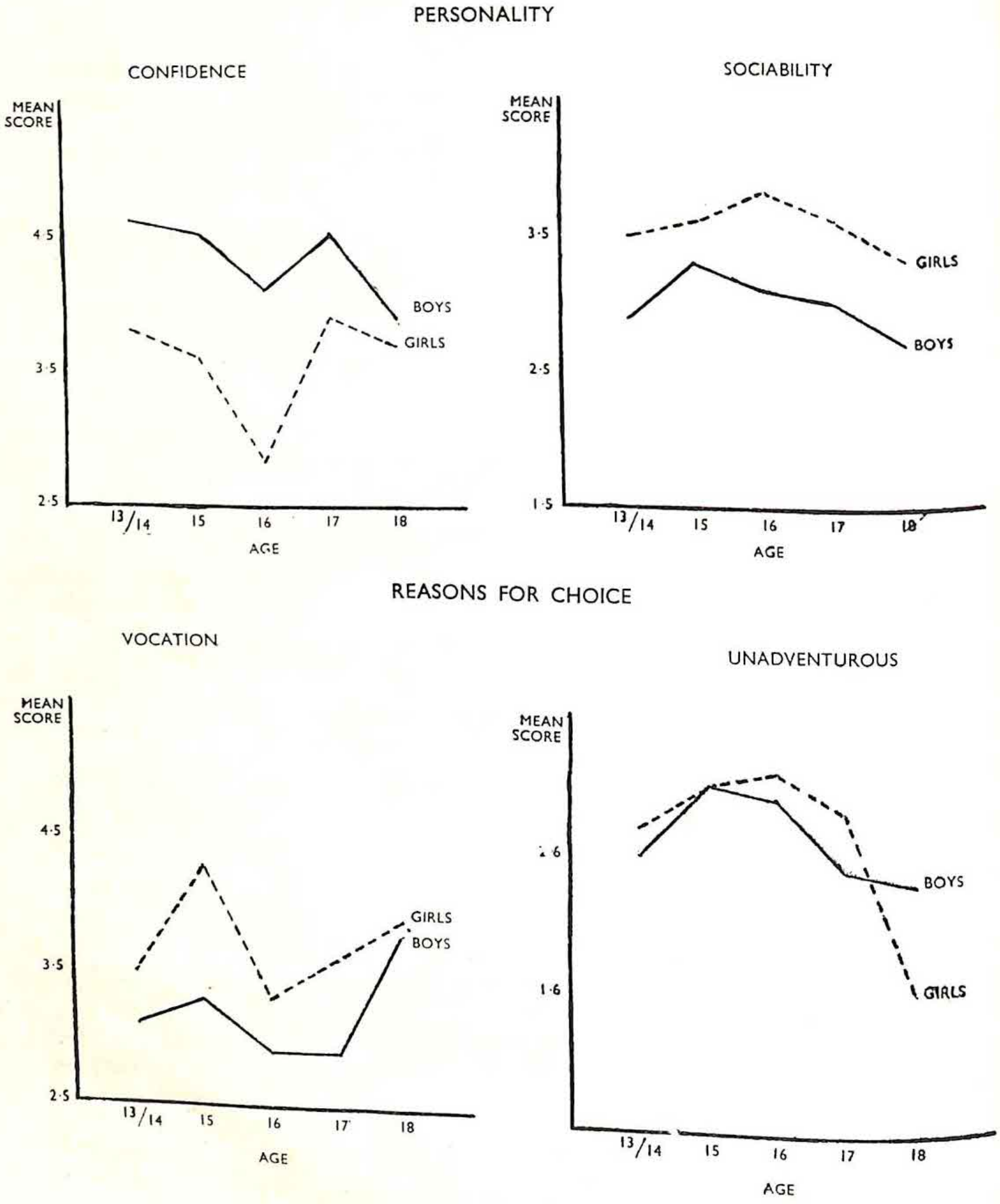


FIGURE 1

Variability of Factor Scores with Age.

equivalent,' and 'University or other choices requiring further training.' For boys the classes were 'Office or office equivalent,' 'University or teaching' and 'Engineering or the armed forces.' If the numbers had allowed it, the groups would have been defined more narrowly. For example, 'Office equivalent' would not have been equated with office work at all; as it stands it must be regarded as including a variety of choices not requiring professional training.

For each factor, the sample was divided into thirds as nearly as the distributions of the factor scores would allow, to give a 3 by 3 table for the boys, and a 3 by 2 table for the girls. The distribution of occupational choices in relation to the factor scores was then tested, using chi-squared.

It seemed possible that any relationship between factor scores and choice might be established later than at the age of the younger subjects. Therefore, another analysis was carried out, considering only those who left school within a year of completing the questionnaire. In order to have expected frequencies of at least five, it was usually necessary to divide this smaller sample only into upper and lower scorers on each factor.

The personality factors appear to influence these subjects in their present choice of jobs only marginally; the extent of the association between low scores on 'confidence' with choice of office work for girls, and high scores on 'sociability' with office work among boy 'leavers,' suggests that in larger groups with more homogeneous choices the relationship between these factors and choice might be shown up more definitely.

Both 'vocational' and 'unadventurous' reasons for choice appear in this sample to be related to job-choice-at-present. The relationship occurs in the expected direction. Those choosing office work score higher on the 'unadventurous' factor, and lower on the 'vocational' factor than those choosing other types of job.

TABLE 6

RELATIONSHIP BETWEEN FACTOR SCORES AND PRESENT CHOICE OF JOB.

		Girls		Boys	
		Whole sample	'leavers'	Whole sample	'leavers'
Personality	Confidence	$P < {\cdot}05$ (scoring categories combined)	—	—	—
	Sociability	—	—	—	$P < {\cdot}05$ (scoring categories combined)
Reasons for choice	Vocation	$P < {\cdot}05$	—	$P < {\cdot}01$	—*
	Unadventurousness	$P < {\cdot}001$	$P < {\cdot}02$	$P < {\cdot}05$	—*

* Because the expected frequencies were less than 5, chi-squared was not calculated.

Thus, the 'reasons' which the subjects think have been important influences in their choice reflect to a significant extent their present choice of job. These reasons, though they may have been formed through many influences do relate back to one of the personality factors ('confidence'), which may, therefore, indirectly have helped to form the choice.

VI.—DISCUSSION.

At the moment, children are for the most part left to form their own opinions about themselves and the type of work they are most fitted to do. In their general talks, Youth Employment Officers try to name the qualities which are required in order to make a success of various occupations. However, Youth Employment Officers play a late and comparatively uninfluential part in occupational choice among grammar school children (Chown, 1958). It is left to the children themselves to discover whether they are suited to any given occupational choice. The measures of 'interests,' 'intelligence,' 'personality traits,' and 'reasons for choice of occupation' studied in the present research were all related to the general type of choice made. It seems, that, even with the very small amount of help they receive, children do consider these factors when making their choice.

It has been said elsewhere that external influences, such as parents and the school, play a large part in determining and consolidating occupational choices : yet children also take into account their own temperaments, apparently almost without realising it. This finding suggests that with added help they would pay even more attention to 'internal' factors. 'Personality' and 'spare time interests' have been shown to fluctuate less with the passing of time than might have been expected in this adolescent group, but even if they had altered greatly, the need for guidance in self-assessment would still be important. It has been suggested by some educationists that introspection is 'bad' for children. It is not, however, that they do not introspect at the moment, but that they have no means of comparing themselves with other people or of judging the occupational significance of what they find.

Very little objective evidence is yet available by which to assess the personality requirements for different jobs. The present investigation highlights the need for further analyses of jobs and occupational success in these terms. It also illustrates the need for vocational guidance in grammar schools to be concentrated on the person as well as the work itself.

VII.—REFERENCES.

BERDIE, R. F. (1943). Factors associated with vocational interest. *J. Educ. Psychol.*, 34, 257-277.

BRAINARD, P. P., and STEWART, F. J. (1932). *Specific Interest Inventories.* New York : Psychological Corporation.

CATTELL, R. B. (1943). Personality : basic traits resolved into clusters. *J. Abnorm. and Soc. Psychol.*, 38, 476-506.

CATTELL, R. B. (1946). *Description and Measurement of Personality.* New York : World Book Co.

CATTELL, R. B., DAY, M., and MEALAND, J. (1956). Occupational profiles on the 16-personality factor questionnaire. *Occup. Psychol.*, 30, 10-19.

CHESIRE, L., SAFFIR, M., and THURSTONE, L. L. (1933). *Computing diagrams for the Tetrachoric Correlation Coefficient.* Chicago : University of Chicago Press.

CHOWN, S. M. (1957). *The Formation of Occupational Choice Among Grammar School Children.* Unpub. Ph.D. thesis, University of Liverpool.

CHOWN, S. M. (1958). The formation of occupational choice among grammar school pupils. *Occup. Psychol.* In press.

EVANS, K. M. (1952). A study of attitude towards teaching as a career. *Brit. J. Educ. Psychol.*, 22, 63-69.

FLEEGE, W. H., and MALONE, H. J. (1946). Motivation in occupational choice among junior and senior high school students. *J. Educ. Psychol.*, 37, 77-86.

GINZBERG, E., GINZBERG, S. W., AXELROD, S., and HERMA, J. L. (1951). *Occupational Choice : An Approach to General Theory*. London : Oxford University Press.

JAHODA, G. (1952). Job attitudes and job choice among secondary school leavers. *Occup. Psychol.*, 26, 125-140, and 206-224.

KUDER, G. F. (1939). *Kuder Preference Recvrd*. Chicago : Science Research Associates.

RITCHIE, F. M. (1930). An enquiry as to the reasons for the choice of occupations among secondary school pupils. *Forum of Education*, 8, 81-90.

STRONG, E. K. (1943). *Vocational Interests of Men and Women*. Stanford, Calif. : Stanford University Press.

VALENTINE, C. W., and RITCHIE, F. M. (1927). Reasons for choice of occupation among secondary school pupils. *Forum of Education*, 5, 85-101.

VERNON, M. D. (1937 and 1938). The drives which determine the choice of career. *Brit. J. Educ. Psychol.*, 7, 302-316, 8, 1-14.;

THE EXTENT AND FACTORIAL STRUCTURE OF FUNCTION FLUCTUATION IN A MIXED GROUP OF ADOLESCENTS AND PRE-ADOLESCENTS

By CHARLES C. ANDERSON

(*Department of Educational Psychology, University of Alberta.*)

Summary. 1.—The extent of fluctuation generally characteristic of a mixed group of pre-adolescents and adolescents was significantly greater than that of pre-scholarship primary children, particularly in Arithmetic. Of all tests, one of Food Aversions showed most extensive evidence of fluctuation.

2.—The difference between the average group index of fluctuation in cognitive and in non-cognitive tests was significant at the 1 per cent. level whereas with the younger groups it had been either insignificant or significant only at the 5 per cent. level.

3.—The size of the group indices of fluctuation for all ten pairs of testing occasions was in phase neither between functions and within groups, nor within functions and between groups.

4.—No general factor was observed in the inter-correlations between individual indices of fluctuation on all six functions.

5.—Further investigation will be necessary before conclusions can be drawn about the effect of the sex factor on contributions to fluctuation in performance during pre-adolescence and adolescence.

I.—Introduction.

'Function fluctuation' is a term orginally coined by Thouless (1936) to refer to genuine quantitative changes in the effectiveness with which a function operates on different testing occasions. Tests of the extent of fluctuation in pre-scholarship classes were recently made by Anderson (1958). So far no one has extended these measures to the secondary school, although an attempt to detect and characterise the fluctuator and to determine the impact of the pre-adolescent's and adolescent's reputed restlessness and moodiness on his cognitive performance must be of significance to all educators who are interested in the pupil's general behaviour as well as in his specific reactions in the learning situation.

The present paper is, therefore, an attempt to extend the application of measures of the extent of fluctuation to the performance of the secondary school child. The general aim is to study developmental evidence concerning fluctuation and thereby to facilitate the production of psychological hypotheses about its factorial structure and psychological correlates, and about the conditions under which it is active. Specifically, the paper has two main aims: firstly, to examine the extent of fluctuation in a selected group of pre-adolescents and adolescents. Clearly we should know more about the nature of fluctuation if these subjects showed more extensive evidence of fluctuation than the child in the primary school. Secondly, the paper aims to examine the factorial structure of the intercorrelations between independent indices of fluctuation derived for each individual and measuring the degree to which the individual is contributing to the amount of function fluctuation observed between any two testing occasions. Basically, this is an investigation into the possibility of deriving for each individual a measure of 'general fluctuation' as a criterion with which other variables might be correlated. The consequent multiple correlation might yield information about the correlates of fluctuation.

II.—EXPERIMENTAL GROUPS AND TESTS.

The experimental group consisted of 58 children, 27 girls and 31 boys, from a rural secondary modern school in England. They were in their second year, and their ages ranged from 11 years 3 months to 13 years 3 months, the average being 12 years 4 months. The highest I.Q. was 122, the lowest 72, on the first testing occasion, and the complete distribution is listed below.

TABLE 1

DISTRIBUTION OF I.Q.'s*

Range	f
120—124	1
115—119	4
110—114	8
105—109	6
100—104	15
95—99	10
90—94	5
85—89	4
80—84	3
75—79	1
70—74	1

N=58
M=100·53
S.D.=10·6

Six sets of parallel tests were administered. The three cognitive tests were the Moray House Intelligence Tests (Adv.) 7 and 8; the Moray House English Tests 16 and 17, and the Moray House Arithmetic Tests 16 and 17. The three non-cognitive tests were the Checking test (sub-test 2 of the Clerical Tests 1 and 2 published by the N.F.E.R.) and tests of Food Aversions and Annoyances, all of which had been given before to the younger pre-scholarship age-group and were chosen to maximise the amount of fluctuation observed in the function. We might presume that certain tests—say, of religious or political attitudes—would be unlikely to elicit much short-term fluctuation, but this is scarcely true of measures of temporary and evanescent moods and feelings. Accordingly, the subjects were asked to rate the strength of their annoyance† in certain specific situations, since most feelings of annoyance are transient and pushed out of the focus of consciousness very soon. The same might be said of aversion to food in general and to certain specific foods in particular, and a simple test of Food Aversions was adapted from Eysenck (1952) by adding a five-point rating scale measuring the degree of aversion after each food listed. Gough (1946) and Wallen (1948) have demonstrated the capacity of this test to discriminate between normals and neurotics, and Wallen concludes that high aversion scores are probably related to poor emotional control. Consequentily it would be interesting to determine whether it yielded the most extensive evidence of

* The average I.Q. may seem unusually high for this type of school, with at least five pupils at an intelligence-level which might be considered adaquate for grammar-school work. This may be ascribed to the possibility that education in rural schools is less closely aligned than is education in urban schools to the strict necessity of pushing pupils through the scholarship examination.

† Adapted from Bennett and Slater (1945) by adding a five-point rating scale after each source of annoyance.

fluctuation. The Checking test had previously showed much evidence of fluctuation, a fact hardly surprising in view of Mackworth's (1948) demonstration that individuals show considerable crude variability in scoring in tests which demand a consistent concentration of attention. In this research, the Checking test has been classified as non-cognitive because adequate performance on it demands neither constructive thinking nor the ability to see relationships nor the recall of learning. Further, it might plausibly be argued that the ability to concentrate attention on this perceptual level is more a product of non-cognitive factors than of cognitive ability.

The two parallel tests (A and B) of each function were administered in a fixed order for five weeks of the school term between October and November. The tests of Verbal Intelligence were given on Monday, the English on Wednesday and the Arithmetic on Friday. The shorter non-cognitive tests were given on Thursday afternoons. Prizes were awarded for improvements in scoring above the average of all previous tests carried out after the first test occasion, the prizes being distributed in such a way that everybody received at least one prize and nearly everybody the same number of prizes. Care was taken to see that the duller children received as many prizes as the brighter children.

III.—THE ANALYSIS OF TEST RESULTS.

The method of testing for evidence of fluctuation is one of Thouless's in which parallel tests of a function are given on at least two separate testing occasions (A_1, and B_1, A_2 and B_2). Qualitatively evidence of fluctuation will appear when an increase (or decrease) in score between A_1 and A_2 is accompanied by a corresponding increase (or decrease) in score between B_1 and B_2. Such a general change in scoring common to the two tests would provide fairly reliable evidence of a quantitative alteration in performance-level which could not plausibly be ascribed to test error alone. Quantitatively, the corresponding increase (or decrease) in scores can be measured by correlating the two difference-scores, $A_1—A_2$ and $B_1—B_2$. Accordingly, the empirically derived formula for testing for evidence of fluctuation between any two testing occasions is ${}^r(A_1—A_2)(B_2—B_2)$. This can be expanded in the form :

$$ {}^r(A_1 - A_2)\,(B_1 - B_2) = \frac{{}^rA_1B_1 + {}^rA_2B_2 - {}^rA_1B_2 - {}^rB_1A_2}{2\sqrt{(1^r - A_1A_2)(1 - {}^rB_1B_2)}}. $$

An estimate of the amount of fluctuation present between any two testing occasions can be obtained (following Thouless) by dividing this quantity by the average of the within-occasion reliabilities, viz :

$$ \frac{{}^r(A_1 - A_2)(B_1 - B_2)}{\frac{1}{2}({}^rA_1B_1 + {}^rA_2B_2)}. $$

This is a group index of fluctuation.

In the present research, in each of six functions every child made on each of five testing occasions parallel scores, that is A_1 and B_1, A_2 and B_2, A_3 and B_3, A_4 and B_4, and A_5 and B_5. For each function separately these ten scores were intercorrelated, yielding forty-five correlations, and all ten possible difference-scores of the type A_1-A_2 and B_1-B_2, derived from $\frac{5+4}{2}$ possible pairs of testing occasions, were intercorrelated. These intercorrelations were transmuted to the group of indices of fluctuation which form the basis of much of the following results.

IV.—TEST RESULTS : THE EXTENT OF FLUCTUATION PRESENT.

For each of the six functions, the appropriate ten group indices of the extent of fluctuation were averaged and these appear below :

TABLE 2

GROUP INDICES DERIVED FOR VARIOUS FUNCTIONS FROM ALL POSSIBLE COMBINATIONS OF TESTS AND RETESTS ON FIVE TESTING OCCASIONS.

Function	Mean
Verbal Intelligence ..	·1616
English	·2794
Arithmetic	·3622
Checking	·5168
Food Aversions	·9347
Annoyances	·4325

Two comments can be made about this table. Firstly, the mean of the group indices for the non-cognitive tests is ·628, that for the cognitive tests ·267. This difference is significant at the 1 per cent. level. Formerly, the same comparison was insignificant for the pre-scholarship group and significant at the 5 per cent. level only for the scholarship group. Further evidence must be advanced in support of this before it can be accepted as a genuine trend with age. Secondly, the test of Food Aversions shows the most extensive evidence of fluctuation, the difference between its mean and that for its nearest rival, Checking, being significant beyond the 1 per cent. level. These two points are evidence of the fact that where cognitive control is missing, or where habitual attitudes have not been built up by social pressures, and the individual is left to his own devices, extensive fluctuation in behaviour and attitude is the rule.

V.—TEST RESULTS : THE EXTENT OF DIFFERENCES BETWEEN OLDER AND YOUNGER CHILDREN.

The amount of fluctuation observed in the three cognitive functions (Verbal Intelligence, English and Arithmetic) across five testing occasions was available for the pre-scholarship (younger) and secondary modern (older) children. The significance of the differences in amount of fluctuation between the age-groups and among the different functions was tested by analysis of variance in the usual way'

TABLE 3

SIGNIFICANCE OF THE DIFFERENCES BETWEEN AGE GROUPS AND AMONG THREE COGNITIVE FUNCTIONS IN EXTENT OF FLUCTUATION ACROSS FIVE TESTING OCCASIONS.

Source of Variance	Sum of Squares	d.f.	Mean Square	F
Functions	·121	2	·061	1·30
Age	·241	1	·241	5·13
Functions X Age	·121	2	·061	1·30
Witihn Groups..................	2·556	54	·047	
Total	3·039	59	—	—

The F for Age is significant at the 5 per cent. level, indicating that the older age-group is generally fluctuating much more than the younger group. The greatest increase in this between-age fluctuation is found in Arithmetic, as can be seen from Table 4.

TABLE 4

SIGNIFICANCE OF THE DIFFERENCE BETWEEN OLDER AND YOUNGER GROUPS IN EXTENT OF FLUCTUATION IN THREE FUNCTIONS.

Function	$M_o—M_y$	t
Verbal Intelligence	·0428	·377
English	·0874	1·116
Arithmetic	·2504	2·620

The t of 2·62 is significant at the 5 per cent. level. Since it is well known that temperamental factors have specially adverse effects on performance in Arithmetic, then we may assume that the temporary and transient presence of these factors in pre-adolescence and adolescence is responsible for this increase in fluctuation in Arithmetic with age.

An exact comparison of the size of the discrepancies between the two age-groups in the Checking, Food Aversion and Annoyances tests is impossible because these tests were previously administered only twice and, therefore, only one group index is available for each function in the younger group.

When the ten group indices for each function were ranked in order of size for older and younger groups, the correlations between the rank-orders were ·339 (Verbal Intelligence), – ·037 (English) and ·531 (Arithmetic). All are insignificant. When the criteria for the older group alone were ranked for each of the three functions, the rank-order correlations were : Verbal Intelligence-English= – ·002 ; Verbal Intelligence-Arithmetic= – ·382, and English-Arithmetic + ·394. Again all are insignificant, indicating that the rise and fall of the criteria in size is not in phase either within groups and between functions or between groups and within functions. Clearly fluctuations conform to no set rhythms.

VI.—TEST RESULTS : THE FACTORIAL STRUCTURE OF FUNCTION FLUCTUATION.

An index of fluctuation was established for every pupil in each of the six functions in the following way. His scores on the A tests and on the B tests were graphed separately and two regression lines drawn through the points, one for A scores, one for the B scores. The shortest distance was calculated from each point to its regression line, so that for every individual in each function there were available two sets of five deviation or ‘ fluctuation ’ measures corresponding to the A and B tests. The correlation between these measures was taken as an index of the extent to which each individual pupil was contributing to the amount of fluctuation evident in the relevant function across all testing occasions. Accordingly, every pupil had six of these correlation-indices, one for each function. These indices were multiplied by 10, normalized to a mean of 50 and a standard deviation of 15, and then intercorrelated.

TABLE 5

INTER-CORRELATIONS BETWEEN INDIVIDUAL INDICES OF FLUCTUATION IN SIX FUNCTIONS.

	E.	A.	Ch.	F.A.	Ann.
Verbal Intelligence	— ·080	·168	— ·108	·319	·014
English		·119	— ·081	— ·128	·165
Arithmetic			— ·270	— ·050	— ·194
Checking				— ·242	— ·063
Food Aversions					— ·340

Crude average r= ·018.

There is clearly no evidence of a general factor in this matrix and, therefore, no hope at the moment of finding the correlates of fluctuation by means of multiple correlation. Function fluctuation is an unpredictable variable characteristic of everybody in varying degrees at different times.

This finding is important for the theory of mental measurement. The rationale of measurement in every science is based on the assumption that the errors of measurement are random in their effect. Campbell (1928, pp. 142-3) has said that " Measurement is accurate if it obeys the laws of measurement . . . and if there are none but irregular errors." Now there is no guarantee that errors due to a fluctuation of function are uncorrelated between testing occasions. In fact, Fiske and Rice (1955, p. 218) have insisted without proof that " . . . intra-individual response variability is not random ; it is a lawful phenomenon. The variability of one's individual responses . . . is determined by more or less enduring factors within the individual." This is a crucial assertion which Fiske (1957) has never been able to justify. His recent contemporary evidence, for example, consists of a matrix of intercorrelations of crude variability scores from eleven non-cognitive tests, the median rho being + ·22, which is quite insignificant for his N=29. The intercorrelations of Table 4, together with the data previously cited by Anderson constitute clear evidence that fluctuation is an uncorrelated and random type of error variable.

VII.—TEST RESULTS : THE SEX FACTOR AND FLUCTUATION.

Do girls at this age-level fluctuate more than boys and in what functions ? In an attempt to answer this question, an analysis of variance was made of the individual indices (in their relatively crude state before they were normalised) after a model supplied by Edwards (1950).

TABLE 6

ANALYSIS OF VARIANCE OF INDIVIDUAL INDICES OF FLUCTUATION BY BOYS AND GIRLS IN SIX FUNCTIONS.

Source of Variance	Sum of Squares	d.f.	Mean Square	F.
Sex	575	1	575	·236
Between subjects of same sex	136,341	56	2,435	—
Total between subjects	136,916	57	—	—
Functions	43,887	5	8,775	3·440
Sex x Functions	22,641	5	4,528	1·776
Pooled subjects x Functions	714,000	280	2,550	—
Total within subjects	780,528	290	—	—
Total	917,444	347	—	—

The F for Functions is significant slightly beyond the 1 per cent. level, and this validates previous findings concerning the fact that non-cognitive functions show more extensive fluctuation than cognitive functions. Neither the F for Sex nor that for the interaction Sex × Functions is significant. However, further investigation reveals that the differences in mean individual indices between the sexes are significant in two functions, Verbal Intelligence and Food Aversion.

TABLE 7

MEAN DIFFERENCES BETWEEN THE SEXES IN INDIVIDUAL INDICES OF FLUCTUATION IN SIX FUNCTIONS.*

	Functions					
	V.I.	E.	A ..	Ch.	F.A.	Ann.
Mean differences	41·91	14·84	−16·83	−20·77	−33·1	−5·971
t	2·44	·801	·845	1·305	2·362	·362

These results are inconclusive, especially in view of the small number of pupils involved.

VIII.—DISCUSSION.

In the present research, pre-adolescents and adolescents show a significantly greater amount of fluctuation than children in the primary school. If it is plausible to attribute this finding either to the restlessness or to the variation in available amounts of energy characteristic of these youngsters (Wattenberg, 1955, p. 67), then we may presume that at least one of the origins of fluctuation lies in a temporary breakdown in the efficiency of cerebral control. This may be a produce of transitory and excessive emotional strength or of fatigue. However, it is clear the fluctuation is an elusive variable and that everyone is capable of contributing to it to a different extent on different occasions. There are at least three ways in which more can be learned about it. Firstly, more information must be made available about the extent of its activity in various functions at different age-levels in both boys and girls. This type of information is indispensable for the formulation of hypotheses about fluctuation. Secondly, the effect on the extent of fluctuation of any experimentally introduced variable might be measured. For example, Mackworth discovered that the administration to adult subjects of 10 mg. amphetamine sulphate (Benzedrine) markedly reduced their variability in scoring on a vigilance test which is formally quite similar to the present Checking test. He concluded that, since the initial level of achievement by the subjects who had taken Benzedrine was not significantly greater than the initial level reached by those who had not taken the drug, then the Benzedrine acted to hold off the encroachments of fatigue. Accordingly, if Benzedrine were administered to an experimental group and the extent of fluctuation observed in that group decreased markedly below the extent of fluctuation observed with a comparable group of subjects who had taken merely placebos, then this would favour the hypothesis that the incidence of fluctuation was correlated with the contemporary coincidence of relative fatigue within the

* A positive mean difference means that girls are contributing more to fluctuation than boys.

nervous system. Thirdly, future researchers might concentrate on correlating individual indices of fluctuation in cognitive tests with performance on tests of personality variables administered on the same testing occasion as the cognitive tests.

IX.—REFERENCES.

ANDERSON, C. C. (1958). Function fluctuation. *Brit. J. Psychol. Monogr. Suppl.*, No. XXX.

BENNETT, E., and SLATER, P. (1945). Some tests for the discrimination of neurotic from normal subjects and the psychometric differentiation of neurotic from normal subjects. *Brit. J. Med. Psychol.*, 20, 271-82.

CAMPBELL, N. R. (1928). *An Account of the Principles of Measurement and Calculation.* Longmans, Green.

EDWARDS, A. L. (1950). *Experimental Design in Psychological Research.* New York: Rinehart.

EYSENCK, H. J. (1952). *The Scientific Study of Personality.* Routledge and Kegan Paul.

FISKE, D. W. (1957). The constraints on intra-individual response variability. *Educ. Psychol. Measmt.*, 17, 317-37.

FISKE, D. W., and RICE, L. (1955). Intra-individual response variability. *Psychol. Bull.*, 52, 217-50.

GOUGH, H. G. (1946). An additional study of food aversions. *J. Abn. Soc. Psychol.*, 41, 86-88.

MACKWORTH, H. (1948). Researches in the measurement of human performance. *M.R.C. Special Report.* H.M.S.O.

THOULESS, R. H. (1936). Test unreliability and function fluctuation. *Brit. J. Psychol.*, 26, 325-43.

WALLEN, R. (1945). Food aversions in behaviour disorders. *J. Consult. Psychol.*, 12, 310-12.

WATTENBERG, W. W. (1955). *The Adolescent Years.* New York: Harcourt, Brace.

AN EXPERIMENTAL INVESTIGATION INTO THE RELATIVE MERITS OF LISTENING AND READING COMPREHENSION FOR BOYS AND GIRLS OF PRIMARY SCHOOL AGE

By W. H. KING*

(*University of London Institute of Education.*)

SUMMARY. 1.—The scores of primary school children on comprehension tests presented orally and visually were compared. Five different types of material were used.

2.—The correlations of the auditory and visual tests with each other and with an Intelligence test were of the order 0·8.

3.—Differences in the mean scores of boys and girls, between boys and girls on the two methods of presentation, and between boys and girls at five levels of intelligence were examined.

4.—There were variations between boys' and girls' results, and between oral and visual presentation ; there was a marked tendency for boys to do better than girls on some of the oral tests with practical and scientific content.

5.—Passages chosen for similarity of content were not necessarily equivalent to the children.

6.—One thing clearly emerges from this research : a single mark for one comprehension test is not a valid or reliable mark for an assessment of comprehension in general ; a number of short passages with variety in content might be a more useful guide.

I.—INTRODUCTION.

A GREAT deal of research has been carried out in the teaching and assessment of school subjects to the neglect, perhaps, of the fundamental sources of learning. " In our society, reading and listening constitute the basic tools of learning as well as the prime media of social intercourse. In the fulfilment of these roles, the importance of reading has never been questioned " (Goldstein, 1940). The process of communication is composed of four skills—the receptive, i.e., listening and reading, and the expressive, i.e., speaking and writing. Yet, although 1,951 studies in reading were reported up to 1939 in the Encyclopedia of Educational Research, there were only fourteen studies related to classroom listening comprehension (Nichols, 1948) ; this, in spite of the fact that in school, and later, more time is devoted to learning by listening than to any other of the language skills. Wilt (1950) reported that children are expected to listen 57·5 per cent. of the time they spend in the classroom and about one-third of this time is spent in listening to the teacher. Rankin (1929) found that the total time devoted to communication by adults was split thus : 45 per cent. in listening, 30 per cent. in speaking, 16 per cent. in reading, and 9 per cent. in writing. There is no evidence that these figures are applicable to our own schools, but they do provide a salutary shock and a very good reason for emphasising accurate recall and reporting of speech. Obviously, any serious lack of proficiency in any of these skills must be a great handicap in the learning situation.

---

* My thanks are due to the following teachers (members of the Teachers' Research Group) who administered the tests : Mrs. L. C. S. Ager, F. G. Bird, E. R. Brown, Miss J. P. Denison, H. C. Deverson, S. Flann, S. G. Melhuish, F. J. Dryden Taylor, C. V. S. Williams.

The main problems are (1) the relative efficiency of visual and oral presentation (learning from printed work or speech), (2) factors contributing to success at the latter. There have been studies at elementary, secondary modern, college and adult levels, though researches among adults have chiefly been concerned with broadcast talks.

II.—OUTLINE OF PREVIOUS RESEARCH.

The results reported vary with the type of material used in the experiment, the degree of selection exercised, age, intellectual level and background experience and the degree of statistical control employed. Carver (1935), using group comparison with equivalent sets of material, found that the relative effectiveness of visual over oral presentation varies directly with the difficulty of the material, and that the effectiveness of listening is greater when the material is simpler; auditory presentation is markedly inferior when strange and unusual material is used. Up to a point, the higher the cultural level, the greater is the capacity to profit from auditory presentation. Finally, he reports that a fairly clear result of this experimental work " is that adults of the lower cultural level profit more by reading than they do by hearing. A good educational background quite naturally gives people an advantage in mastering both printed and spoken words, but oddly enough it gives them relatively greater advantage with the spoken word."

Carver emphasises the differences between reading and listening comprehension. Printed words are " experienced more as related items in larger groupings than in isolated units." The number of words grasped by a single perceptual grouping will depend on the difficulty of the passage and the familiarity of the reader with the material The reader progresses at his own speed, grouping words and phrases as he wishes; going back, if necessary, to study particular words and contents.

In listening when words are separated in time, the listener will have great difficulty in making his own groupings, and such groupings are more often determined by the pauses and intonation of the speaker than the free choice of the listener. He also makes the point that untrained minds are unable to listen intelligently to material which they could perhaps with difficulty understand by reading.

In an experiment using four narrative, five expository and seven descriptive passages for listening comprehension, M. K. Hollow (1955) found significant correlations between listening scores and intelligence (0·42), and reading comprehension (0·55). Corey (1934) examined college freshmen on recall and found that immediate recall is better for materials students have read than for the same materials heard in lectures, and that there was " a tendency for students scoring in the highest psychological test quartile to do relatively better on reading than on lecture tests."

Larsen and Feder (1940) found that, with college freshmen, a group low in scholastic aptitude was about equal in reading and oral comprehension, and the median group showed some superiority in reading comprehension though the difference was not statistically significant. The top group did show a superiority in reading comprehension and " students high in scholastic aptitude comprehend difficult material better by reading than by hearing, and easy material almost equally well by any method. Students of low ability do slightly but not significantly better when listening to easy materials." They found that the correlation between reading and listening comprehension for this group was 0·62± ·04. With American undergraduates, Goldstein (1940), in a study of reading and listening comprehension at different rates, obtained a correlation of 0·78 between the results of the two types of comprehension ; listening comprehension was generally superior to reading and markedly so with easier passages. In the detailed study by Nichols (1948), excerpts from college lectures were given to 200 college freshmen. He found that listening comprehension involves a number of factors not always present in reading comprehension, including sex (men superior), intelligence, correct English usage, size of vocabulary, interest, physical fatigue of listener and audibility of reader.

Schonell (1935) in an experiment with senior school children, obtained a correlation of 0·42±0·05 between intelligence scores and immediate memory tests with broadcast talks ; similarly work with adults has been reported by Vernon (1950) and Belson (1952). One thing stands out clearly from the published results of these researches ; lisening comprehension is a complex skill and tests of vocabulary, English Usage, recall, intelligence and so on only give particular facets and in no way add up to the complete ' gestalt.'

Possibly the most extensive research on listening was that carried out by Brown (1949) in the preparation of the Brown-Carlsen Listening Comprehension Test (1953). Correlations with various intelligence tests were of the order 0·7. A split half reliability coefficient of 0·85 is claimed for this test and it apparently gives higher correlations with college achievement than does reading comprehension. Several American achievement batteries (e.g., " Evaluation and Adjustment Series," World Book Co.) make use of sub-tests of auditory comprehension.

III.—PLAN OF THE INVESTIGATION.

In the present research, an attempt was made to compare the scores of boys and girls on tests of visual and auditory comprehension and to relate these scores to the results of a verbal intelligence test and four key tests described below. After discussion with a group of teachers interested in educational research, ten passages, each of about 100 words, were put together : (a, f) descriptive, (b, g) conversational, (c, h) imaginative, (d, i) practical, (e, j) scientific. These passages were cyclostyled on one side of quarto paper ; the questions were on the reverse side so that they were answered without further reference to the passage. One passage and relevant questions are given in the appendix.

The five remaining tests of the battery were : (1) a verbal intelligence test of eighty-six questions (I), (2) a test of ability to follow directions (D), (3) a test of immediate recall using numbers, letters and words (R), (4) a test of word meanings (W), (5) a test of attention (N).

The test of following directions (D) contained twenty questions. Five of these questions were based on the five letters A to E, and in one the children were asked to write down the series obtained by reversing the middle three letters. Other questions were based on the combination of numbers 1 to 8 and letters A to H. In the test of attention (N), the children were given answer sheets on which were spaces for the answers to twenty questions. Three digits were read and the children were asked to place a tick against the question number only if the first number was the largest and the second, the smallest. All these tests, except the Intelligence test, were presented orally and were checked for time on a group who did not take part in the main experiment.

The Intelligence test was given to 475 boys and girls in their last term at nine primary schools in the London area. On the results the schools were divided into two groups (A and B), approximately equal in numbers and mean and range of intelligence (Table 1).

TABLE 1

Group	N	M	σ
A	236	45·5	20·6
B	239	45·3	21·7
Boys	225	44·8	21·9
Girls	250	45·8	20·6

Administrative details were discussed with the teachers and notes on the administration (speed of reading, answer time, etc.) were circulated. Importance of normal classroom presentation was stressed. The ten comprehension tests (a—j) were given on alternate mornings as far as possible, so that complete results on the fifteen tests were not available until a month after the start of the experiment.

Group A completed the comprehension passages (a—e) orally (i.e., the passages were read to them and they wrote their answers to the questions on a given form) and the passages (f—j) visually (i.e., the children read the passage, turned over their paper and answered the questions on the other side). Three minutes were allowed for the presentation of the passage by either method and four minutes for answering the questions. Group B reversed the order of presentation and took (a—e) visually and (f—j) orally. On the results of the intelligence test the scores of the ten tests were obtained for five levels of intelligence (Table 2). An investigation into the sigfinicance of differences between the means of boys and girls at the five levels of intelligence, and between the two forms of presentation of the comprehension passages, completed this stage of the investigation.

TABLE 2

Average Scores for each of the Ten Tests at each of the Five Levels of Intelligence from Totals of Boys' and Girls' Scores on all Comprehension Tests.

Mean of 3,000 scores for ten tests of comprehension =9.361.						
	I_1	I_2	I_3	I_4	If	Mean scores for each test
a	11·68	11·25	10·62	9·00	5·88	9·69
b	11·25	11·00	10·12	9·07	6·43	9·57
c	13·98	13·25	12·38	10·25	7·88	11·55
d	10·22	9·32	8·40	6·28	3·35	7·51
e	11·75	10·51	9·02	5·63	2·47	7·88
f	12·45	11·23	10·58	9·20	5·40	9·77
g	11·63	11·12	10·65	10·03	6·63	10·01
h	13·15	12·73	12·45	11·08	8·02	11·49
i	12·65	11·58	9·40	6·95	3·88	8·89
j	10·87	8·82	7·40	5·52	3·62	7·24

IV.—Results.

(*a*) *Analysis of Variance.*

Disproportionate frequencies in the cells prevented a complete analysis of variance of the 475 sets of comprehension scores, but a five factor analysis was carried out on 60 sets of scores chosen at random from each grade of intelligence, i.e., 300 sets in all or 3,000 scores. The five variables were (i) method of response (T), auditory or visual, (ii) order of presentation (P), auditory tests followed by visual or the reverse, (iii) sex (S), (iv) intelligence (I), (v) contents of ten paragraphs (A).

The I and A factors were both significant at the ·001 level and the T factor at ·05 level, but the P and S factors were not significant. There was evidence, however, since the S × A interaction was significant at the ·001 level, that boys and girls do react differently to the same material.

A second analysis of variance, omitting order of presentation (P) again showed that the S factor was not significant, but the method of response (T) was significant at 1 per cent. level and the I and A factors were highly significant. To present a clearer picture of the interactions, a three factor analysis (T, S, A) was calculated for each of the five levels of intelligence. At all five levels the A factor (variation in contents of the five pairs of passages) was significant at the 0·1 per cent. level. In the lowest intelligence group the T and S factors were significant at the 1 per cent. level and the T × A interaction was significant at the ·01 per cent. level, indicating that the relative superiority of visual or auditory presentation among children at this level varies with the type of passage. In the highest grade of intelligence, A × S was significant at the 5 per cent. level, showing that boys and girls reacted differently to different types of passages.

(*b*) *Regression Coefficients.*

All the scores on the auditory comprehension tests and the visual comprehension tests were separately summed (^{c}aud, ^{c}vis), and the inter-correlations of these totals with the scores on the intelligence test (I) and the four key tests, directions (D), recall (R), word meanings (W) and attention (N) were calculated. Test N has the smallest set of correlations with the other tests. The correlations of ^{c}aud and ^{c}vis (0·8219) is the highest recorded, but both ^{c}aud and ^{c}vis have correlations of the order 0·8 with intelligence.

Table 3 gives the regression coefficients obtained from the correlation matrix using in turn ^{c}vis and ^{c}aud as criterion. The multiple correlation of the battery with ^{c}aud is 0·837 and with ^{c}vis is 0·829. With the auditory comprehension tests, intelligence contributes 38 per cent. of the variance, word meaning 22 per cent., and ability to follow directions 10 per cent. With the visual comprehension tests, intelligence accounts for 50 per cent. of the variance, word meaning 13 per cent. and recall nearly 8 per cent. As a check on the mechanical accuracy of these computations, Doolittle's and Aitken's methods were both used.

TABLE 3

REGRESSION COEFFICIENTS.

	Criterion : ^{c}aud				Criterion : ^{c}vis			
	β	σ	Sig.	% Variance	β	σ	Sig.	% Variance
I	0·4778	·058	Yes	38	0·6197	·044	Yes	50
N	0·0049	·037	No	0	−0·0482	·028	No	0
R	−0·0077	·047	No	0	0·1258	·036	Yes	8
W	0·3028	·048	Yes	22	0·1895	·036	Yes	13
D	0·1474	·051	?	10	−0·0124	·039	No	0

Multiple R = 0·837. Multiple R = 0·829.

(c) Comparison of mean scores of boys and girls in the ten comprehension tests.

(i) *Differences between boys on oral and visual tests.*—Boys' mean scores were higher on the auditory tests than on the visual tests in eight out of the ten passages. Three differences were significant.

(ii) *Differences between girls on oral and visual tests.*—On five of the passages girls obtained higher mean scores on the visually presented tests (c, e, f, i, j) and on the remaining five higher mean scores were obtained on the orally presented tests (a, b, d, g, h). Only one of these differences was significant at the ·05 level.

(iii) *Differences between boys at five levels of intelligence.*—Because overall means masked possible differences (shown in the three factor variance analyses) mean scores on auditory and visual tests were examined at each level of intelligence. Of the fifty pairs of means examined, the boys gained higher mean scores on thirty-nine of the auditory tests. No pattern was discernible between the I levels, though the results suggest boys tend to do better than girls on the auditory presented tests.

(iv) *Difference between boys and girls means at five levels of intelligence.*—Table 4 shows the main differences in the approach of the children to the ten comprehension passages. Boys tend to return higher mean scores on the oral tests, and girls on the visual test. Though there are variations in the tests in each category there are several where the same differences appear in at least three out of the five levels of intelligence and there is some evidence that tests d, i (practical) are understood better by boys and tests a, f, c, h (descriptive and imaginative) by girls when presented visually. When the tests are presented orally, boys return higher mean scores on d, e, i, j (practical and scientific) but nowhere is the division at all clear cut.

TABLE 4

SIGNIFICANT DIFFERENCES BETWEEN SCORES OF BOYS AND GIRLS AT DIFFERENT LEVELS OF INTELLIGENCE.

				Higher mean scores returned by			
Intelligence groups	Significant difference in test	Level of Significance	Sex	Boys in visual tests	Girls in visual tests	Boys in oral tests	Girls in oral tests
I_1	a (visual) i (auditory)	·01 ·05	Girls Boys	b, d, e, i	a, c, f, g, h, j	a, d, e, h, i, i	b. c. f. g
I_2	a (visual) i (auditory)	·02 ·01	Girls Boys	d, g, i, j	a, b, c, e, f, h	b, e, f, h, i, j	a, c, d, g
I_3	i (auditory) d (auditory)	·05 ·05	Boys Boys	b, d, h	a, c, e, f, g, j	b, c, d, e, h, i, j	a, f, g
I_4	i (auditory) e (auditory)	·01 ·05	Boys Boys	b, d, e, f, g, i, j	a, c, h	b, c, d, e, f, g, i	a, h, j
I_5	—	—	—	d, i	a, b, c, e, f. g. h. j	b, f, i, j	a, c, d, e, g, h

(*v*) *Differences between girls at five levels of intelligence.*—There was less difference between the scores on auditory and visually presented tests. The girls gained higher mean scores on the visual presentation on twenty-six of the fifty tests.

(vi) *Differences between all boys and girls on the auditory tests.*—Boys gained higher mean scores on five of the ten tests, but for only one (passage i) was this difference significant (·01 level). The girls did better on the descriptive and imaginative passages, the boys on practical and scientific.

(vii) *Differences between all boys and girls on the visual tests.*—Here girls obtained higher mean marks on eight of the ten tests; significantly so on passage a (·01 level) and passage g (·05 level). The boys did better on the two practical tests (d, i) as they did when these two tests were presented orally.

(viii) *Differences between all boys and girls on all tests* (*visually and orally presented*).—Table 5 gives the means of total scores for boys and girls on both methods of presentation. There was no significant difference between the responses of boys and girls whether the test was given orally or visually; girls did not react more to one or other of the methods of presentation, but the boys returned significantly higher scores (·01 level) on the auditory presentation.

TABLE 5

10 Tests	Boys Mean	Girls Mean	
auditory......	9·79	9·61	N.S.
visual	9·24	9·53	N.S.
Significance ..	·01	N.S.	—

V.—Appendix.

It was an ordinary little house, in an ordinary little street, and it belonged to a man who followed the very ordinary trade of tanning leather for boots and shoes or anything else that is made of ordinary leather. Just notice that nobody said it was an ordinary man that owned the house. A man is one thing and his trade is another. The most important part of Joseph's house was the cellar, a two-roomed cellar. The right-hand room was full of round and square ditches filled with strong-smelling tan and the left-hand room was full of deep vats where the leather soaked, while the river flowed right up to the house wall itself to do the soaking.

Above the cellar was Joseph's workroom, full of his big wooden tools and above that again were two little rooms, side by side, and a little back room. That was all—a little ordinary poor house.

1.—How many rooms in the house?
2.—How many rooms in the cellar?
3.—What did the man do for a living?
4.—What was the name of the man?
5.—Name the two things described as ordinary? (i) ....
(ii) ....
6.—What was the most important part of the house?
7.—What is mentioned that is not ordinary?
8.—What was contained in the left-hand cellar?
9.—The wall of the house was .... because along one side of the house was a ....

VI.—REFERENCES.

BELSON, W. A. (1952). *The Comprehensibility of 'Topic for Tonight.'* B.B.C. Listener Research Report.

BROWN, J. I. (1949). The construction of a diagnostic test of listening comprehension. *J. Exp. Educ.*, 18, 139-146.

BROWN, J. I., and CARLSEN, G. R. *Listening Comprehension Test.* New York: World Book Co.

CARVER, M. E. (1935). *The Psychology of Radio.* Ch. IX. (Editors: H. Cantril and G. W. Allport.) New York: Harper.

COREY, S. M. (1934). Learning from lecturer v. learning from reading. *J. Exp. Psychol.*, 19, 245-264.

GOLDSTEIN, H. (1940). *Reading and Listening Comprehension at Various Controlled Rates.* New York: Teachers' College, Columbia University, Contribution to Education No. 821.

HOLLOW, SISTER MARY (1955). Listening comprehension at the intermediate grade level. *Elementary School Journal*, Vol. 56, No. 4, Dec., 1955.

LARSEN, R. P., and FEDER, D. D. (1940). Common and differential factors in reading and hearing comprehension. *J. Exp. Educ.*, 31, 241-252.

NICHOLS, R. G. (1948). Factors in listening comprehension. *Speech Monographs*, Vol. XV, No. 2, pp. 154-163. Iowa.

RANKIN, P. T. (1929). Listening ability, its importance, measurement and development. *Chicago Schools Journal*, 12, 177-179.

SCHONELL, F. J. (1935). Broadcasting in the senior school. *Brit. J. Educ. Psychol.*, 5 299-322.

VERNON, P. E. (1950). *An Investigation into the Intelligibility of Educational Broadcasts.* B.B.C. Listener Research Report.

WILT, M. E. (1950). A study of teacher awareness of listening as a factor in elementary education. *J. Educ. Res.*, 43, 626-636.

THE INFLUENCE OF RELIGIOUS EDUCATION UPON THE DEVELOPMENT OF CHILDREN'S MORAL IDEAS

By F. H. HILLIARD

(*University of London Institute of Education.*)

Summary. Up to the beginning of adolescence the majority of children believe that God upholds the moral law or ideal by rewarding good behaviour. At the same time, the majority have already either rejected or begun to wonder about earlier and cruder ideas about God's punishment of wrong-doing. As a result of the development of the ability to reason critically, an increasing number of adolescents question, and in many cases abandon, these earlier ideas about God rewarding the good and punishing the bad. Some adolescents and adults, however, retain the basic belief that moral attitudes and decisions should be determined in the light of religious beliefs. The majority of adolescents and adults believe, more vaguely, that religion can and should help them to live a morally good life, though on what grounds this belief is maintained cannot be stated with any certainty. Not all who hold this view believe that religion is *essential* for living a morally good life.

I.—Introduction.

(1) *Definitions.*—(*a*) Religious education here refers mainly to Christian religious education, and religious instruction means Christian religious instruction based largely on the Bible. It is recognised that much religious education exists beyond these limits which cannot be considered in an article which is both restricted in length and also concerned with a particular set of facts and problems.

(*b*) Religious education here includes not only religious instruction, but all religious ideas and attitudes, implicit and explicit, which may influence children from several directions, of which the chief are the home, the school, Sunday School and the Church.

(2) *Religious education part of a complex influence.*—Any attempt to isolate the influence of religious education from that total complex influence which society brings to bear upon the moral development of the child cannot wholly succeed and the present attempt must be regarded in the light of this fact. The influence of a child's 'peer-group,' of the moral 'atmosphere' of his schools, of literature, the film, radio, and television, are all more or less potent factors in the field of moral, as of other aspects of a child's development. With the implementation of the provisions of the Education Act of 1944, however, religious education has become a statutory task of the schools. The vast majority of children, therefore, are receiving some religious education, and a not inconsiderable number also attend Sunday Schools—at any rate up to about the time at which they transfer to a secondary school. On the face of things, therefore, religious education seems to be a potentially important influence among all those which affect the development of children's moral ideas.

(3) *The scope of this survey.*—Religious education could be expected to influence the growth of children's moral ideals principally in two ways. It could help to shape their ideas about the importance of specific virtues and ideals by bringing such virtues and ideals before them in stories and teaching contained in the Bible, in the history of the Church, in hymns and prayers, and in religious teaching generally. It could also introduce the child to the idea that what are commonly called divine 'sanctions'—the will of God, heaven, hell, etc.—underlie and re-inforce the whole moral order of things. It is with the

second of these two kinds of influence that this article deals, partly on account of limitations of space, but mainly because, of the two, that of religious 'sanctions' upon the development of children's moral ideas seems to the present writer more urgently to need some clarification.

(4) *Sources.*—Apart from the use made of published works and research noted in the list of References, the article has drawn upon the results of an investigation which the present writer initiated. It is referred to below as 'Q.T.T.S.' and details are given both in the text and in an Appendix. It should be made clear that since this was a 'pilot' investigation, information provided by it has been treated as having only tentative value.

II.—THE ASSOCIATION OF MORAL IDEAS WITH RELIGION IN CHILDHOOD AND EARLY ADOLESCENCE.

It seems clear that the young child connects ideas of right and wrong, good and bad, with the idea of God and regards God as rewarding good behaviour and punishing wrong-doing. Some would argue that the child assumes this association long before he is given formal religious teaching. It also seems plain that with many children this attitude continues up to about the stage of early adolescence.

Bovet, writing in 1925, noted the great variety of religious experience observable among young children and found its roots in a religious sentiment which he believed to be an extension and modification of the filial sentiment. He suggested that the child has a view of its parents which clothes them with the divine attributes of classical theology—omnipotence, omniscience, and moral perfection. Later, as the religion of parents is abandoned under the stresses and strains of the intellectual crises of early childhood, there dawns for the child the religion of the heavenly Father. He maintained that the attribute of moral perfection with which the child imbues his parents is one of the most persistent features of this early attitude.

Piaget quotes Bovet's views on this matter with full approval. He maintains that the gravity of the crisis provoked by the discovery of a fault, and particularly an injustice, in the parents, furnishes convincing proof of the contention that the young child endows his parents with the attribute of moral perfection. He refers to an instance of a child who, accused and punished in error, ended by convincing himself that he was guilty of the fault with which he was charged. He comes to the further conclusion, moreover, that though the child's education can provide him with the idea of the perfect God which allows the necessary transference to take place, there is the danger that too early an insistence on divine perfection may result in setting up God as a rival to the parents in the child's mind.

We have very little knowledge of the religious ideas of pre-adolescent children, but such information as is available lends some support to the idea that belief in God at this stage continues to include as an important element the belief that He is concerned with upholding what is right and good. Indeed, even at the stage of early adolescence many children appear to find no difficulty in accepting certain naïve notions of the manner in which God upholds right as against wrong, by rewarding good behaviour and punishing wrong-doing.

Kuhlen and Arnold (1944) reported the results of an investigation into the religious beliefs of children between the ages of 12 and 18. A total of 547 children (257 boys, 290 girls) completed a questionnaire in which the subject was asked

Bureau Ednl. Psy. Research
DAVID HA [illegible] AINING COLLEGE

to indicate whether he 'believed', 'did not believe,' or did not know whether he believed or not but 'wondered about,' a number of statements representing various religious beliefs. The subjects included 174 children of age 12, 243 of age 15, and 130 of age 18. Almost all the children in all three groups had affiliations with Protestant or Roman Catholic Churches; only a very small number were Jews or had no affiliation with religious bodies at all. Three of the statements and the responses made to them were as follows:

TABLE 1

EXTRACT FROM KUHLEN AND ARNOLD'S FINDINGS.

Statement	Age	% Who "Believe"	% Who "Do Not Believe"	% Who "Wonder About"
1. Only good people go to heaven.	12 15 18	72 45 33	15 27 32	13 27 34
2. God is someone who watches over you to see that you behave yourself, and who punishes you if you are not good.	12 15 18	70 49 33	18 37 48	11 13 18
3. Hell is a place where you are punished for your sins.	12 15 18	70 16 13	16 21 30	13 27 34

If we consider for the moment only the responses made by subjects in the 12-year-old group, it becomes apparent that approximately 70 per cent. of those who responded accepted these rather crude interpretations of a basic belief in God as the upholder of the sanctity of the moral law or ideal.

Bradshaw reported (1950) findings which tend in the main to support those of Kuhlen and Arnold, though with some differences. He attempted a study of the religious beliefs of 298 children between the ages of 11 and 18 (181 boys and 117 girls). Their distribution according to age was as follows:

11—12·11 years	13—14·11 years	15—16·11 years	17—19 years.
72	86	85	55

They were all receiving religious instruction in their schools. About half of the boys had also attended Sunday School up to the age of 14, and some for a longer period. A larger proportion of the girls than the boys had attended Sunday School, and generally for a longer period. Part of this investigation was conducted by means of a questionnaire which invited the subject to indicate whether he regarded statements representing certain religious beliefs as 'true,' or 'false,' or ones about which he 'did not know' whether they were true or false.

Three of the statements which are closely similar to those selected from Kuhlen and Arnold's investigation received responses as follows:

TABLE 2

EXTRACT FROM BRADSHAW'S FINDINGS.

Statement	Age	% Who said "True"	% "False"	% "Do not know"
1. Good people go to heaven.	11—12·11	68	14	18
	13—14·11	54	19	27
	15—16·11	45	20	35
	17—19	31	38	31
1. God will punish you if you are bad.	11—12·11	45	40	15
	13—14·11	38	49	13
	15—16·11	38	35	27
	17—19	24	50	26
3. Bad people go to hell.	11—12·11	33	34	33
	13—14·11	19	43	38
	15—16·11	15	39	46
	17—19	13	50	37

If we again confine our attention for the moment to the responses made in the 11—12·11 years group and compare them with those of Kuhlen and Arnold's 12-year-old group, it will be noticed that :

(1) The implied belief underlying statement (1) that God rewards good conduct is supported by just under 70 per cent. of the children in each case and disapproved of by about 15 per cent.

(2) The percentage of those who agree with the cruder ideas about God punishing the wicked, and the wicked going to hell, in statements (2) and (3) is lower in Bradshaw's findings than in those of Kuhlen and Arnold (45 per cent. and 33 per cent. respectively, against 70 per cent.). It is worth noting, however, that almost 50 per cent. of the children here are prepared, up to the age of 13, to accept the idea that God upholds the moral order of things by punishing the wicked. It will also be seen that even the very naïve idea behind statement (3) is actually rejected by not more than 34 per cent. of the subjects.

The responses made to four questions put in 'Q.T.T.S.' to a total of 220 students in one university department of education and three training colleges are also relevant here :

TABLE 3

FIRST EXTRACT FROM "Q.T.T.S." FINDINGS.

Question No.	% answering "Yes"	% "No"	% "Not Sure"
5. Were you taught that God required people to be good ?	85	9	4
6. Have you ever been taught that God would punish you if you did wrong ?	49	44	5
7. Was the idea of Heaven presented to you as an incentive to being good ?	54	40	5
8. Was the idea of Hell presented as a deterrent to being bad ?	42	51	4

It will be observed that the forms of these questions were necessarily different in certain respects from those used in the other investigations referred to above. Here also, allowance must be made for the fact that the subject would not always have been able to recollect precisely what earlier attitudes he had about these matters. Nevertheless, the general impression left on the minds of 85 per cent. of the subjects is that in childhood God was presented to them as the upholder of the moral law or ideal. A considerably smaller number (54 per cent., 49 per cent., 42 per cent) could. recall having had presented to them the type of religio-moral ' sanctions ' implied in Questions 7, 6 and 8, though even here the numbers are either just above or not far below the 50 per cent. level.

Broadly speaking, these findings suggest that at the beginning of adolescence, while the majority of children have rejected some of the earlier ideas about God's ' punishment ' of wrong-doing, the majority also continue to believe that God upholds the moral law or ideal. It is impossible to say, in general, whether this last-mentioned belief is the outcome of influences which proceed from one direction more than another. The basic attitude implied in it is one of the oldest in the history of human thought : it lies at the root of the primitive notion of ' taboo,' which carries within itself the germ of the later idea, found in one form or another in all the great religions, that religion and morality are intimately associated.* It is an attitude which is certainly inculcated in the minds of many young children by their parents, who use it as a means of reinforcing early teaching on behaviour. The bed-time prayers repeated by young children after their mother still frequently end with some such petition as " . . . and make me a good boy (girl). Amen." There is little doubt, moreover, that a good deal of the school and Sunday School teaching given to young children, not only in actual lessons but also through the hymns and prayers used, reflects the same basic attitude—that God requires them to be good, and sometimes (though the indications are that most teachers and parents now avoid the suggestion) that God directly punishes wrong-doing.

III.—CHANGES IN IDEAS DURING ADOLESCENCE.

The evidence that is avilable suggests that during adolescence the earlier ideas that God rewards the good and punishes the bad are increasingly questioned and often abandoned. Both Kuhlen and Arnold's and Bradley's findings above reveal a marked decrease, as adolescence proceeds, in the numbers of those who formerly accepted more primitive ideas about the divine punishment of wrong-doing. They also reveal a decrease in the numbers of those who formerly accepted the much more basic idea that the good life earns a divine reward. Replies received in ' Q.T.T.S.' suggest further confirmation of these tendencies. Question 9 asked, with reference to the ideas contained in Questions 5, 6, 7 and 8 quoted above, " Did you later question the ideas referred to in 5, 6, 7 and 8? If so, (*a*) which ? and (*b*) at approximately what age ? " 150 subjects answered as follows :

---

* " In taboo we touch the borderlines of religion and morality. Breach of taboo is the outstanding, indeed, almost the sole, crime of rudimentary societies . . . Taboo may not be, as we judge matters, specifically moral, but . . . at least it presents a moral aspect." WATERHOUSE, E. S. (1936). *The Dawn of Religion*, p. 70. London : Epworth.

TABLE 4

SECOND EXTRACT FROM " Q.T.T.S." FINDINGS.

Belief referred to in Question No.	Was questioned at approximate age—							
	Before 11,	11,	12,	13,	14,	15,	16+	years by
5	0	0	3	12	12	12	19	subjects
6	5	3	9	14	22	23	30	,,
7	9	4	12	9	21	27	28	,,
8	8	4	13	9	25	26	27	,,

These figures suggest that children and young people engage in the most intense questioning of these earlier ideas during the periods of middle and later adolesccence.

There is, of course, nothing surprising in the fact of these questionings and changes of attitude. Burt has suggested that destructive or critical reasoning (as distinct from inductive reasoning), normally begins to function at about the age of 11 years and the first manifestations of doubt appear to follow the development of this capacity. The questioning and in many cases the consequent abandonment by many adolescents of earlier ideas which associate God in naïve fashion with the fact and the maintenance of the moral law or ideal, is an aspect of adolescent religio-moral development which must be expected and regarded as normal.

At the same time it is important to notice that the fact that adolescents often reject these earlier ideas does not necessarily imply that they also abandon the basic conviction that the moral law or ideal is somehow associated with the existence of God. There are marked indications that many adolescents and adults consider that belief in God and religious faith generally can and should help people to live morally good lives. Moreton reported (1944) the results of an enquiry into the attitudes to religion of 354 adolescents and 414 adults. The adults (258 men and 156 women) replied to a questionnare in which they registered agreement or disagreement, strong agreement or strong disagreement, or neither agreement nor disagreement, with a number of statements. Two statements and the average scores which they received are relevant here :

TABLE 5

EXTRACT FROM MORETON'S FINDINGS.

Statement No.	Statement	Score
1	Religion makes people live better lives	+ ·37
21	The highest ways of living are those taught by the great religious leaders	+ ·62
	(Maximum positive score= +2 ; maximum negative score= −2.)	

Forrester reported (1946) the results of an investigation into the attitudes of adolescents towards, *inter alia*, their own spiritual development. As part of the investigation, she tested 346 children (157 boys, 189 girls) between the ages of 13 and 17+, by means of a questionnare on religion. The subjects were asked to indicate their agreement or disagreement with a number of statements dealing with the nature and purpose of religion. The following statements and the scores they received may be noted :

TABLE 6

EXTRACT FROM FORRESTER'S FINDINGS.

Statement No.	Scores at Age				
	13	14	15	16	Sex
1. The chief business of religion is to teach people how to be good.	·18 ·93	·46 ·90	·19 ·72	·27 ·05	B G
4. People do not need religion in order to be good.	·45 – ·48	·29 ·05	1·27 ·04	·54 ·60	B G
6. The right kind of behaviour really depends upon our having the right kind of belief.	·45 1·03	·29 ·75	– ·12 – ·12	– ·23 – ·20	B G
11. Some of our Scripture lessons should deal with present social evils and problems.	1·09 1·24	1·46 1·45	1·50 1·60	1·30 1·40	B G
13. Our responsibilities to society have nothing to do with religion.	– ·18 – ·48	– ·33 – ·70	·15 – ·88	·45 – ·52	B G

It will be observed that a majority of these adolescents appeared to believe that religion teaches people how to be good, while at the same time also believing that people could be good without it. They disagreed, however, as to whether religious belief had a direct bearing upon conduct. With some exceptions in the case of the boys of 15/16 the majority assume that religious beliefs may have a bearing upon social morality. Miss Forrester herself suggests that her findings as a whole indicate that the adolescent thinks that religion should be practical and should teach people how to be good. She notes, however, that more than half the children, especially the older children, think that people do not necessarily need religion in order to be good.

Similar attitudes are recorded in the replies made to five questions asked in 'Q.T.T.S.'

It seems, therefore, that a not inconsiderable number of adolescents and adults retain from earlier days a conviction, often hard to define, that religion can and should assist in the maintenance and development of moral ideas and conduct. Whether this conviction springs ultimately from the vague persistence of that age-old attitude noted above which from earlier times led to the association of religion and morality ; or from a respect for the moral strength of the religious teaching of, e.g., Christ ; or from a religious principle such as is involved e.g., in the New Testament dictum that he who loves God must love his brother also, is bound to depend in the last resort upon the existence or otherwise of religious convictions in the individual, and if they exist, upon their quality and depth. Obviously, therefore, it is not a matter about which generalisations could

TABLE 7

Third Extract from " Q.T.T.S. Findings.

Question No.	% answering " Yes "	% " No "	% " Not Sure "
12. Do you believe that moral values derive their authority from purely human sources, e.g., from the need for individuals to live harmoniously in society ?	17	73	8
13. Do you believe that the moral teaching of Jesus can provide a satisfactory guide in matters of :			
(*a*) private morality (e.g., honesty in private life)	93	5	0
(*b*) social morality (e.g., family relationships)	90	7	0
(*c*) international morality (e.g., treatment of enemies)	80	12	5
20. Do you consider that belief in God can help young people to live a good life ?	93	2	3
21. Do you consider that belief in God is *essential* if young people are to live this kind of life ?	38	50	9
22. Do you consider that moral education can best be given :			
(*a*) solely through religious education ?	13	85	0
(*b*) partly through religious education ?	85	13	0
(*c*) separately from religious education ?	21*	75*	2*

be made. That certain young people, however, believe that their moral ideas and attitudes are influenced by their religious convictions, which in turn must be to some extent the result of religious education, is clear. ' Q.T.T.S.' provided some relevant information. Question 11 asked : " Have you maintained them (i.e., earlier ideas mentioned in Questions 5-8) in a different form, e.g., in a belief that we should " be perfect as our Heavenly Father is perfect ? " and received replies as follows :

% answering " Yes "	% " No "	% " Do Not Know "
63	27	5

In a group of whom 75 per cent. regarded themselves as church-goers the figure of 63 per cent. is low ; this may be partly due to the fact that a number of the subjects said that they did not understand the meaning of the question.

Of another group of 399 training college students (73 men, 326 women) 157 (23 men, 134 women) said that the most important reason for their resisting temptation to behave in a dishonest, unkind, or selfish way was because they wanted to live in accordance with the principles of their religion.* In this particular group, therefore, 39 per cent. indicate that they would expect their religious convictions to have a direct bearing upon their moral ideas and attitudes.

---

* Some who answered " Yes " to (*b*), thought that moral education could also be given effectively apart from religious education.

It is possible that if in this last group those whose religious education had helped to shape their moral ideas and attitudes in less clearly-defined ways were added to the total of 157, the final figure would be considerably higher. We here find ourselves confronted in fact with the difficulty of being unable to say with any finality where ' religious ' influences end and ' non-religious ' influences begin. In a society in which Christian influences have been at work for fifteen centuries and more, what was once ' religious ' may now appear simply ' democratic ' or ' humanistic,' and it is difficult to see how any investigation could provide any final picture of the full influence of religion upon moral ideas and attitudes in contemporary English society.

IV.—CONCLUSIONS.

The most important general conclusion to be drawn from this survey would seem to be that while the majority of adolescents experience considerable dissatisfaction in regard to earlier ideas and teaching about the divine ' sanctions' of morality, they continue to look to religious education and religion itself to assist them in the development and maintenance of their moral ideas and standards of conduct. This being so, it follows that religious education which deals effectively, in middle and later adolescence particularly, with religio-moral principles and their application to current problems is likely to meet with an interested response from adolescents.

V.—APPENDIX.

DETAILS OF "Q.T.T.S."

The questionnaire was completed by 220 students in one university department of education and three training colleges (approximately 30 men and 190 women).

Their age-range was from 18-22 years.

75 per cent. were church-goers and 25 per cent. non-church-goers.

10 per cent. had never attended Sunday School: 90 per cent. attended for varying periods. 98 per cent. had received some religious instruction in school, and 97 per cent. in both primary and secondary schools.

The questionnaire was divided into two sections:

Section A comprised six questions dealing with general details—religious education, and church affiliation.

Section B comprised twenty-three questions dealing with the subject's attitudes to and/or recollections of:

Bearing of religion upon morality (1-4).
Teaching received on religious ' sanctions ' (5-8).
Present attitude to religious ' sanctions ' (9-15).
Effect of religious education upon moral ideas (16-19).
Present attitude to value of religious education for children's moral education (20-23).

---

* I am indebted to Mr. James Hemming for this information, supplied in a private communication.

VI.—REFERENCES.

BOVET, P. (1925). *Le Sentiment Religieux et la Psychologie de l'Enfant.* Engl. Trans. (1928), *The Child's Religion.* London : George Green.

BRADSHAW, J. (1950). *A Psychological Study of the Development of Religious Beliefs in Children and Young Persons.* M.Sc. Thesis, London University.

BURT, C. (1931). The mental development of children, essay in. *Board of Education Report on the Primary School,* 42-3, and Appendix III. London : H.M.S.O.

FORRESTER, J. F. (1946). *A Study of the Attitudes of Adolescents Towards their own Intellectual Social and Spiritual Development.* Ph.D. Thesis, London University.

KUHLEN, R. G., and ARNOLD, M. (1944). Age differences in religious beliefs and problems during adolescence. *J. Genet. Psychol.,* 65, 291-300.

MORETON, F. E. (1944). Attitudes to religion among adolescents and adults. *Brit. J. Educ. Psychol.,* 14, 69-79.

PIAGET, J. (1929). *The Child's Conception of the World.* Engl. Tr. London : Kegan Paul.

—— (1932). *The Language and Thought of the Child.* Rev. Engl. Tr. London : Kegan Paul.

—— (1932). *The Moral Judgment of the Child.* Engl. Tr. London : Kegan Paul.

THE BEHAVIOUR OF IDIOTS INTERPRETED BY PIAGET'S THEORY OF SENSORI-MOTOR DEVELOPMENT

By MARY WOODWARD

(*Senior Psychologist, Fountain Hospital, London*)

SUMMARY. Piaget's observations on sensori-motor intelligence and the concept of permanent objects in the development of the normal infant were applied to the study of 147 mentally defective children. The postulated order of difficulty of responses typical of each of Piaget's sensori-motor stages was confirmed. Repetitive behaviour with objects was readily classifiable into the types of manipulation which Piaget terms secondary, derived secondary and tertiary circular reactions ; most of the hand mannerisms resembled the hand movements of infants to which Piaget has given the name of primary circular reactions. The postulated correspondence between the type of problem-solving or grasping behaviour and the type of circular reaction was found in 43 per cent of the cases. The postulated correspondence between the stage of sensori motor intelligence and object concept development was found in 87 per cent of the cases. It was concluded that the sensori-motor activities of severe mental defectives could be classified into the six main types distinguished by Piaget.

I.—INTRODUCTION.

THE behaviour of severely mentally defective children appears, at first sight, to be purposeless and unlike that of the normal child of any age ; their repetitive hand mannerisms in particular contribute to this impression. It was, however, noticed during clinical observations that these mannerisms resembled the infant's hand movements to which Piaget (1953) has ascribed a developmental function. It followed from this observation that the significance of idiot behaviour might be illuminated by reference to Piaget's observations and concepts. It was, therefore, postulated that the sensori-motor activities of severe mental defectives, directed to their own bodies or to objects, are not a set of purposeless behaviour patterns specific to pathological types of mental deficiency ; on the contrary they may be classified into the six main types of sensori-motor activity which Piaget (1953, 1955) has distinguished in the development of the normal infant during the first two years.

The period of sensori-motor intelligence is, in Piaget's (1950) scheme, the first in a sequence of stages which culminate in abstract thinking. Inhelder (1944) has already used the later stages of intuitive thinking and concrete operations to classify moderate cases of mental deficiency, but her studies did not extend down to the sensori-motor stage. Since Berlyne (1957) has recently made a comprehensive summary of Piaget's whole scheme, only the first stage will be briefly described here. This was derived from extremely detailed observations of three infants during the first two years of their lives. The aspect of the theory which is termed sensori-motor intelligence, deals with the modification of elementary behaviour patterns into more complex ones adapted to the solution of simple problems (Piaget, 1953). The other aspect concerns the development of the child's knowledge of the objective world as consisting of a spatial field of permanent objects whose movements have causal interactions and occur in a temporal sequence (Piaget, 1955). Time did not allow the study of causality and the spatial and temporal fields in the present inquiry, which dealt with sensori-motor intelligence, and the development of a concept of a permanent object, since experiments could more readily be devised for these two aspects. The stages for each of them are given below :

Sensori-motor intelligence.

Stage I—Extension of reflex activity.

Stage II—Integration of reflex activities with higher activities : co-ordination of hand with mouth, vision with hearing and vision with grasping, by repetition of simple behaviour patterns (primary circular reactions) directed to the child's own body and not acting on objects (e.g., hand movements).

Stage III—Reproduction of chance-made movements (secondary circular reactions) that produced an effect on objects which the child had seen or heard (e.g., the child makes objects swing or rattle).

Stage IV—Application of existing behaviour patterns to new situations (e.g., the child knocks down an obstacle that prevents him from grasping an object).

Stage V—Discovery of new behaviour patterns in new situations through experimentation : (*a*) when the child handles objects, instead of the simple repetition of stage III, there is variation in the effect, apparently directed at exploring the object ; (*b*) new means are discovered to solve a problem (e.g., the child obtains an object that is out of his reach by pulling an extension of it).

Stage VI—Invention of new means to solve a problem by foresight (e.g., the child uses a stick to obtain an object that is out of reach).

Object concept development.

Stages I and II—The child makes no response when objects go out of his perceptual field.

Stage III—The child continues a movement that is in progress when the object goes out of his perceptual field, but does not search for it.

Stage IV—The child briefly searches for an object that goes out of sight, but confers an absolute position on it by searching in the first place where he found it, despite having seen it disappear elsewhere.

Stage V—As stage IV except that the child now takes account of changes of position he observes, and searches briefly in the correct place.

Stage VI—The child searches for a longer period of time in a manner that suggests awareness of the existence of unperceived objects.

The applicability of this system of classification to mental defectives was examined by the following hypotheses :

I.—(*a*) That the order of responses when ranked according to the decreasing number of defectives who show them will agree with the order of their appearance observed by Piaget in normal infants at successive ages ; (*b*) that mental defectives who solve problems which are classified by Piaget at a given stage of sensori-motor intelligence will show responses of earlier stages.

II.—The type of circular reaction shown in manipulative play or hand movements will correspond with the type of grasping or problem solving behaviour in mental defectives as in normal infants.

III.—The stage of sensori-motor intelligence and the stage of object concept development will correspond in mental defectives as in normal infants.

II.—SUBJECTS.

The subjects were resident at a hospital for severely mentally defective children of many different clinical types. The total populations of narrow age ranges were preferred to a sampling of all ages. Two contrasting age groups were chosen, 14 to 16 years (the oldest available) and 7 to 9 years. Group A, which was studied in greater detail, consisted of all children in these age groups who failed to achieve a basal age of two years on the Terman-Merrill scale and who had no motor or sensory disability. In order to confine the group to children who had consistently functioned as idiots, two who had previously shown a higher level of function were excluded. This gave 65 idiots to form Group A. All but 12 (who were in the younger age range) could walk alone ; only one child used word combinations appropriately, and five others had some echolalic speech ; 56 appeared to have no verbal comprehension.

Less detailed information was obtained from a second group (B) of 82 children when they were examined for ordinary clinical purposes at admission or subsequently. Group B is thus more selected. It contained fewer emotionally unstable children (28 per cent. compared with 38 per cent. in Group A), but this difference is not statistically significant ($\chi^2=1{\cdot}8$, $p.>{\cdot}05$, d.f.l.). Children were described as emotionally unstable if they cried or whimpered on entering a strange room, if they displayed marked visual avoidance, began a mannerism when approached, ran away or struck an offered hand. Group B ranged in age from 7 months to 15 years, the median age being 6·2 years. Some of the younger children were not idiots, but it was considered of interest to include imbecile infants. Sixty-five children in Group B, like Group A, were unable to succeed with all the first six items of the Binet scale ; the median Binet I.Q. of the remaining 17 children was 33.

The clinical findings and classifications for both groups are given in Table 1.

TABLE 1

CLINICAL FINDINGS AND CLASSIFICATIONS.

	Group A	Group B
Mongolism	9	30
Phenylketonuria	5	9
Tuberose sclerosis	3	2
Microcephaly	3	4
Von Recklinghausen's disease	2	0
Icterus gravis neonatorum	2	1
Birth injury	1	2
Congenital syphilis	1	0
Post vaccinal encephalitis	1	0
Spina bifida	0	3
Galactosuria	0	1
Cretinism	0	1
Hypertelorism	0	1
Acrocephaly	0	1
Unclasssified, epileptic	19	4
Unclassified, not epileptic	19	23
Totals	65	82

III.—METHOD.

Children in Group A were observed for at least two sessions, children in Group B only once.

Mannerisms and manipulation of play material were observed during the child's spontaneous activities with toys which were presented separately in the same sequence for all children ; those offered first were still available later. Strings that could be pulled to make objects swing and rattle were already fixed when the child came in. Bricks and nests of cups and tins were presented so that the child could see them. A bell, rattles and a rubber doll that squeaked were sounded when held outside the child's visual field. The investigator was passive unless the child wished to involve her in his play. All observations were recorded at the time.

If the child made no movement towards the toys, the grasping sequence described by Piaget was repeatedly operated in reverse. The toy was placed (*a*) so that the child could see it and his hand together, (*b*) in his mouth, (*c*) touching his hand and (*d*) in his palm.

Six problems were devised, two from each of Piaget's last three sensori-motor stages, as follows :

Stage	*Problem.*	*Method.*
IV	Transparent (T) Screen and Opaque (O) screen.	The screens were placed between the child and an object he was about to grasp. The solution was to remove the screen.
V	Support.	An object was shown to the child and then placed out of his reach : (*a*) On a piece of cardboard which was in his reach.
V	String.	(*b*) Attached to a piece of string, with the other end near the child's hand.
VI	Rake.	(*c*) Near a stick shaped like a rake.
VI	Bars.	Sticks and long toys were placed behind the bars of a box in such a position that the child had to turn them through 90 degrees in order to obtain them. They were inserted through a side door, not through the bars, since that would have demonstrated the solution.

Since the children could walk to an object placed out of their reach, a large box with vertical bars at the front was made, and the objects, support, etc., were put inside it. The child could then solve the problem only by using the instrument.

For the object concept experiments, the child's attention was attracted to a sweet or toy which was hidden, while he was still looking, under various covers ; these were part of the play material previously offered, so that the cover itself would not have the attraction of novelty. First the object was half hidden. Another object was then completely hidden under cover A (one of the nest of cups). If the object was found, cup A was left on the table and another object was quickly hidden under cover B (one of the nest of tins). In successive experiments the period of time the object was out of sight was lengthened by increasing the number of covers placed on top of the object. Up to five covers were used. The child had to remove them to obtain the object. Each experiment was repeated, with variation in place and cover, until two unambiguous successes or failures were obtained. These experiments could not be carried out with the eight children who focussed only momentarily on objects, nor with those who were distressed when the object was hidden. In the latter event a toy was hidden unobtrusively while the child was playing.

For the problems and object concept experiments a sweet was used as the experimental object whenever possible, but when the child rejected or ignored sweets, a toy in which he had shown interest was used. Care was taken to avoid assuming that the child had solved the problem when he obtained the object merely as a by-product of an interest in the support, string or cover ; the criterion for success was that the child immediately discarded the instrument or cover and picked up the experimental object. In cases of doubt repreated observations were made in other sessions.

Children in Group B were observed after the analysis of results from Group A. If they succeeded with problems more difficult for Group A, they were not given the easier ones.

IV.—RESULTS.

A.—Order of responses. In Table 2, items from each of Piaget's stages are ranked in decreasing order of frequency, by the number of children in Group A who showed a successful response, defined as follows :

Sucking objects. :	The child sucked objects accidentally contacted, but could not direct them to his mouth.
Co-ordination of :	
vision and hearing.	The child turned his head to sounds.
vision and grasping.	The child was guided by sight when he grasped objects.
Manipulations.	The child repeatedly handled objects to produce an effect on them ; children who put objects in their mouths or held them still to look at them, without changing the position of the object, were not included.
The six problems.	Successful solution as already described.

TABLE 2

ORDER OF SUCCESSFUL RESPONSES.

Responses	Solved by	Stage
Sucks objects	65	I
Directs hand to mouth	65	II
Co-ordination of vision and grasping	62	II
Manipulates objects	53	III
Co-ordination of vision and hearing	49	II
T. screen	37	IV
O. screen	32	IV
String	25	V
Support	22	V
Bars	12	VI
Rake	6	VI

It will be seen from Table 2 that, apart form the co-ordination of vision and hearing, the two orders are the same. The 65 children in Group A were assigned a stage number, which was determined by the point of their first successful response, working backwards from stage VI, e.g., a child who failed the problems of stages V and VI and who solved one stage IV problem would be assigned to stage IV. (In both groups A and B successful solution of one problem of a given stage was significantly related to success with the other problem of the same stage, χ^2 giving a p. beyond ·001 for each of stages IV, V and VI).

Leaving aside the co-ordination of vision and hearing, 59 of the 65 children, when so classified, showed all the responses of all earlier stages. Two children failed one earlier problem and one failed two earlier problems ; three did not manipulate objects. Thirteen children did not turn towards a sound, but ten of them gave an avoidance reaction by turning in the opposite direction, covering up their ears, or beginning a mannerism. It cannot, therefore, be assumed that they had not co-ordinated vision with hearing.

Ability to make a successful response was not always immediately apparent. Six children in Group A and one in Group B grasped objects on sight only after repeated trials. One child used his mouth to pick up objects, and his ability to grasp objects on sight was demonstrated only when he was placed in a play pen and the toy held so that he could reach it with his hand but not with his mouth. Three children grasped on sight only after they had for about half-an-hour been grasping objects under conditions which Piaget observed to precede the co-ordination of vision with grasping in the infant. One child (aged 9 years) who had been wearing gloves to prevent sore fingers from excessive sucking, at first made no move towards objects he saw ; instead he opened his mouth. He did, however, grasp objects when the investigator touched his lips with them, and eventually he grasped objects placed on the table.

The second child (aged 14½ years) at first grasped objects only when they touched her hand, though when they were put on the table within reach, she looked at them attentively, as if desiring them. When they were moved near her hand, she visually followed them, but did not take them. After 15 minutes, however, she grasped a rattle when it and her hand were both in her visual field. After this she still focussed on new objects placed on the table, but made no move to take them, though she did so as soon as they were placed near her hand so that she saw both together. Finally, however, she brought two hands to a rattle on the table and grasped it. The same sequence was subsequently observed in another child aged 3½ years.

B.—Circular Reactions.—The results on the correspondence between problem-solving and grasping behaviour on the one hand and circular reactions on the other are given in Table 3 for both Groups A and B. Manipulations of objects and hand mannerisms which were classifiable as circular reactions, as defined by Piaget, are as follows :

Circular reactions.	*Manipulations or hand movements.*
Primary (Stage II).	Holding up one hand at face level and intently looking at it ; moving one hand, with or without fingers twiddling, watching, or not looking ; clasping both hands together or rubbing them ; playing with fingers ; opening or closing fists.
Secondary (Stage III).	Simple, repetitive banging, shaking, swinging of objects, without variation.
Derived Secondary (Stage IV).	Repeatedly picking up an object to drop it from a height of a few inches ; turning it round in order to examine it.
Tertiary (Stage V).	Spinning or rolling objects, banging them successively on different surfaces and appearing to compare the sound and dropping them from various heights and watching the fall.

Children who displayed more than one type of circular reaction are classified in Table 3 by that which Piaget considered the most complex (e.g., a child showing all types would be classified under tertiary, one showing secondary and primary under secondary), though this was not necessarily the most frequent type. Children showing no circular reactions are shown in column 6 (other manipulations) and column 7 (no manipulations or hand movements). Piaget described no circular reactions for stage VI, so children who succeeded with the rake or bars problems are omitted from Table 3. The 40 children who solved the string or support problem are shown as stage V, while those who failed these problems, but solved one screen problem are shown as stage IV (N=15). Children (N=21) who showed the behaviour of stage III (failing all problems manipulating objects and grasping on sight) are shown separately from children (N=9) who showed the behaviour which marks the end of stage II (grasping on sight). The 12 children who did not grasp on sight have been classified separately.

TABLE 3

CIRCULAR REACTIONS COMPARED WITH PROBLEM-SOLVING OR GRASPING BEHAVIOUR (GROUPS A AND B).

	N in group	Circular reactions				Other manipulations (v)	No activity
		Primary II	Secondary III	Derived secondary IV	Tertiary V		
Did not grasp on sight II beg.	12	**11**	0	0	0	0	1
Grasped on sight II end	9	5	0	0	0	0	4
Handled objects III.	21	0	**19**	2	0	0	0
Screens (T. or O.) IV	15	0	9	**2**	4	0	0
String or support V	40	1	11	7	**10**	8	3
Totals	97	17	39	11	14	8	8

Cases of correspondence are marked in black type in Table 3. If the nine children who grasped on sight but did not manipulate objects are regarded as not showing the expected circular reaction (secondary), there is correspondence in 43 per cent. of the cases (45 per cent. in Group A and 41 per cent. in Group B). Thirty-three non-corresponding cases showed a less complex type of circular reaction than expected, and only six children displayed a more complex type of manipulation than was expected by their grasping or problem-solving behaviour.

The eight children shown in column 6 all solved stage V problems; in spontaneous play they built or aligned objects, or put them in and out of containers. Piaget classifies these activities at stage V in the development of notions of space. Their manipulations thus correspond in this way, although they did not show the expected tertiary circular reactions. Stage V manipulations of one type or the other were shown significantly more often by children who solved stage V problems than by those who failed them ($\chi^2 = 10.08$), p. between ·01 and ·001 for 1 d.f.). Stage IV maniupliations, however, appear to be equally frequent among children solving stage IV problems and those failing them. Among the children who failed all problems, secondary circular reactions were significantly associated with ability to grasp on sight ($\chi^2=11{\cdot}92$, $p.<{\cdot}001$, d.f.1). As expected by Piaget's observations, none of the 12 children who grasped objects only on tactile contact manipulated objects except to suck them or look at them. Hand movements which were classifiable as primary circular reactions (referred to as type A) were an almost continuous activity in 11 of these 12 children; they were an occasional activity of 32 per cent. of the children who grasped on sight, the difference being significant at beyond the ·001 level (χ^2, with Yates' correction applied, being 13·4, d.f.1).

A second type of mannerism, however (type B), was an occasional activity of two of the 12 children who did not grasp on sight, and was more frequently displayed by 25 children who grasped on sight. These children repeatedly shook their hands, jerked their arms or tapped or banged their hands on the table top. It will be seen in Table 4, which compares the two types of mannerism, that only nine of the total 56 manneristic children in Groups A and B exhibited both types. The association of type B with grasping on sight is significant ($p.<{\cdot}05$, $>{\cdot}02$, $\chi^2=4{\cdot}37$ for 1 d.f.). A third type of hand mannerism was shown by two children who repeatedly moved their arms in wide sweeps on the table top.

TABLE 4

MANNERISMS.

Mannerisms	Grasps on		Totals
	Sight	Touch	
Type A only	21	9	30
Type B only	17	0	17
Both types	7	2	9
	45	11	56

C.—Object Concept development.—The stage of object concept development and sensori-motor (s.m.) intelligence are compared in Table 5. Problem-solving and not manipulative activity is used as the basis of classification for s.m. stages IV to VI, as in Table 3. Children who failed all problems and showed secondary circular reactions are classified as stage III. Children who grasped on sight and did not manipulate objects are classified as stage II together with those who grasped by touch only.

Children were classified for stages of object development by the number of covers removed, as follows:

Stage II—None.

Stage III—None, unless the object was only half covered, or the child held and dropped an object, and followed the line of fall.

Stage IV—One (in one place only).

Stage V—One or two (in several places).

Stage VI—Three or more.

In fact, all children classified at Stage VI removed five covers; none removed three or four and then failed with more. Cases of correspondence were 87 per cent. of the total group of 147 (81 per cent. in Group A and 91 per cent. in Group B).

TABLE 5

Comparison of Stages of Sensori-Motor Intelligence and Object Concept Development (Groups A and B).

		Object concept stage					Totals
		II	III	IV	V	VI	
Stage of sensori-motor intelligence	II	**19**	2	0	0	0	21
	III	1	**20**	0	0	0	21
	IV	0	8	**0**	6	1	15
	V	0	0	0	**39**	1	40
	VI	0	0	0	0	**50**	50
Totals		20	30	0	45	52	147

The complete lack of correspondence at stage IV is partly due to the fact that eight children who were classified at s.m. stage IV solved only the transparent screen problem (with the object visible). They were thus unlikely to show the behaviour of object concept stages IV to VI.

D.—Non-correspondence and other factors.—The relation of non-correspondence to epilepsy and emotional instability (as described in section III) was examined, with results as in Table 6.

TABLE 6

Discrepancies in Behaviour Compared With Emotional Instability and Epilepsy

	Circular reactions and other behaviour					S.M. and O.C. stages			
	Discrepancy		Correspondence		Totals	Correspondence	O.C. higher	O.C. lower	Totals
	N	%	N	%	N	N	N	N	N
Emotionally unstable	24	53	10	19	34	38	8	2	48
Stable	21	47	42	81	63	90	2	7	99
Totals	45	100	52	100	97	128	10	9	147
Epileptic	16	35	10	19	26	28	2	5	35
Not epileptic	29	65	42	81	71	100	8	4	112
Totals	45	100	52	100	97	128	10	9	147

Emotional instablity and/or epilepsy was present among 75 per cent. of the children whose manipulation did not correspond as expected with their problem-solving or grasping behaviour. Each condition separately was more common among these children than among the cases of correspondence, but only the difference in respect of emotional instability reached statistical significance (p. of $< \cdot 001$, $\chi^2 12.23$, d.f.1.).

The number of cases of non-correspondence between sensori-motor intelligence and object concept development is too small for statistical analysis.

V.—DISCUSSION.

The postulated agreement between the order of responses (*hypothesis I*) has been confirmed, but problem-solving and grasping behaviour corresponded less often with circular reactions shown in manipulative play or hand movements (43 per cent. of the cases) (hypothesis II) than they did with object concept development (87 per cent. of the cases) (hypothesis III). Two explanations may be offered to account for this. One is that mental defectives do not develop their manipulative and problem-solving ability together in the manner observed by Piaget. The other is that the mentally defective children concerned had a capacity for response that was not exhibited. Responses might be inhibited in emotionally unstable children ; thus, the significant association of emotional instability with non-correspondence supports the latter alternative. Secondly, the comparison between the sensori-motor and object concept stages, at any rate in the later stages, involves the comparison of two more highly motivated situations (solving a problem or removing covers in order to obtain a desired object). whereas spontaneous activity (by which circular reactions were determined) offers no more than its own reward. That some mental defectives have a capacity for response which is not immediately evident, is indicated by the fact that the ability of seven children to grasp on sight was not at once apparent.

Three of these grasped on sight after a period of about half-an-hour during which their grasping followed the sequence described by Piaget as preceding the co-ordination of vision and grasping. Whether these sequences were a revival of movements that had previously led to the co-ordination of vision and grasping or whether they were now emerging for the first time, the behaviour observed in these three cases of severe deficiency follows the same course as that noted by Piaget in his observations on normal infants. However, since these children ranged in age from 3½ to 14½ years, it is unlikely that this co-ordination of vision with grasping had never occurred before. And if so, some explanation is required to account for the fact that it should subsequently lapse. (Only one child was epileptic.) Two explanations may be suggested. One arises from Hebb's (1949) distinction between infantile learning as slower and more permanent and learning at older ages as less stable. His view that early learning is more lasting because later learning tends to utilise it instead of replacing it, would suggest that the learning of the co-ordination of vision and grasping was less permanent in the three mentally defective children than in the normal child because there was little subsequent learning which might utilise it. The amount of other learning is presumably least in the most backward ; secondly, the children concerned, particularly the child wearing gloves, had few opportunities for grasping new objects. Alternatively, other consequences of the very slow rate of development may have been more important : within a few months of acquiring ability to grasp on sight the normal infant has developed more complex behaviour patterns. A mental defective developing at one-tenth the normal rate may take

2½ years to achieve further differentiation or combination of existing behaviour patterns. This would involve the repetition of behaviour in unchanged form for this period, and under these conditions the behaviour pattern may disappear. This may perhaps be likened to the role of repeated or prolonged presentation of the same stimulus in the elimination of a conditioned response.

The prominent features of mannerisms is not the non-appearance of a behaviour pattern, but its prolonged retention. The interpretation offered for type A (likened to Piaget's primary circular reactions) is suggested by Piaget's observation that the infant's prolonged visual interest in his hands drew attention to their grasping function. In his view it is the simultaneous vision of hand and object, first of the object in his hand and then beside it, which leads the child eventually to co-ordinate vision with grasping. These activities at this period of infant development have also been observed by Valentine (1950) and Bühler (1930). Since nearly all the mentally defective children who had not achieved this co-ordination displayed a similar visual interest in their hands, and there was a significant decrease in the frequency of these mannerisms among children who had co-ordinated vision and grasping, it seems likely that the movements have the same function as in the infant. The bizarre quality of idiot mannerisms to which Earl (1934) drew attention, is possibly due to the uneven development of mental defectives in different functions, which Binet (1916) pointed out. The older idiot who has the sensori-motor development of an infant of four months appears qualitatively different from a baby because of his greater size and more advanced locomotor development (walking).

The other main groups of hand mannerisms, which did not resemble primary circular reactions, occurred chiefly among children who could grasp on sight. They were similar to the movements the child makes when he handles objects after he has begun to grasp on sight. Accordingly, it is suggested that these mannerisms are combinations of movements which are developed when the child handles objects in a simple repetitive manner, and are then continued in the absence of objects. Support for this view is provided by the observation of a child of 11 months: having dropped a rattle she had been shaking, she continued the shaking movement with her hands.

VI.—CONCLUSIONS.

It has been found in this enquiry that sensori-motor responses ranked in the order in which mental defectives show them, in decreasing order of frequency, agree closely with those observed by Piaget to appear in normal infants at successive ages. Furthermore, most of the idiots showing the responses of Piaget's later stages also exhibited those of earlier stages. This suggests that the sensori-motor development of severe mental defectives follows the sequence described by Piaget. The correspondences that have been found suggest that the level reached by a given mentally defective child in his sensori-motor activities, indicated by problem-solving or grasping, usually corresponds with the stage he has reached in the development of a concept of permanent objects, and, at any rate with the stable child, with his spontaneous activities. Thus, it seems reasonable to conclude that the sensori-motor activities of pathological mental defectives may be classified into six main types similar to those observed by Piaget in normal infants.

Further research could usefully extend this sytem of classification to grossly unstable and physically handicapped children: it offers a means of interpreting the child's spontaneous behaviour, when this is all that is available for observation; secondly, the method the child uses to attempt to solve a problem is as informative as a successful solution. Practical applications are at present limited by the fact that Piaget's observations were made on only three children, and by the consequent need for research along these lines with larger numbers of normal children. However, some of the items of behaviour which constitute Piaget's stages have been studied in larger groups (e.g., by Bühler, 1930; Shirley, 1931; and Gesell, 1940). Single items of behaviour, such as finding a sweet under a cup, are of course of little value when considered in isolation. They have, however, more significance when viewed as a point reached by the child in the developmental sequence of one function which develops in relation with other functions. Regarded in this light, any one item of behaviour by which Piaget characterises a stage would be interpreted as a single item in the child's total pattern of sensori-motor behaviour; and furthermore, this total pattern of behaviour would be regarded as a distinctive stage in a complex developmental process which follows the same intelligible sequence in nearly every child.

VI.—REFERENCES.

BERLYNE, D. E. (1957). Recent developments in Piaget's work. *Brit. J. Educ. Psychol.*, 27, 1-12.

BINET, A., and SIMON, T. (1916). *The Intelligence of the Feeble-minded.* Vineland: New Jersey.

BÜCHLER, C. (1930). *The First Year of Life.* New York: John Day.

EARL, C. J. C. (1934). The primitive catatonic psychosis of idiocy. *Brit. J. Med. Psychol.*, 14, 230-253.

GESELL, A. (1940). *The First Five Years of Life.* New York: Harper.

HEBB, D. O. (1949). *The Organisation of Behaviour.* New York: Wiley.

INHELDER, B. (1944). *Le diagnostic du raisonement chez les Debiles Mentaux.* Neuchâtel: Delachaux Néistlé.

PIAGET, J. (1950). *The Psychology of Intelligence.* London: Routledge and Kegan Paul.

—— 1953). *The Origins of Intelligence in the Child.* London: Routledge and Kegan Paul.

—— (1955). *The Construction of Reality in the Child.* London: Routledge and Kegan Paul.

SHIRLEY, M. M. (1931). *The First Two Years.* University of Minnesota Press.

VALENTINE, C. W. (1946). *The Psychology of Early Childhood.* (3rd edition.) London: Methuen.

RESEARCH NOTE.

Mental Age Scales

By H. MADDOX
(Department of Education, University of Birmingham.)

I.—Introductory.

The results of most intelligence tests are now expressed as normalised standard scores. Both the Moray House and the National Foundation for Educational Research (NFER) tests are so constructed that the standard deviation of I.Qs. or standardised scores is 15 points, a score of 100 corresponding to the 50th percentile. 115 and 85 to the 84th and 16th percentiles. This system at least ensures consistency in the meaning of I.Qs. or standardised scores from one age group to another.

Many schools, however, still use the older verbal group tests, such as the Simplex Junior and the Essential Junior Intelligence test, for internal school examinations. These tests yield so-called classical I.Qs. based on the ratio of mental age to chronological age. In these older tests the mean I.Q. and the standard deviation of I.Qs. is usually not given for particular age groups. Such scales have so often been criticised by statisticians as being incapable of interpretation, that their continued use is a commentary on the conservatism of test publishers and teachers in this country. Most head masters have learnt from experience that these tests yield inflated I.Qs., particularly for brighter children, but continue to use them for practice purposes and as a general guide, without taking the figures too literally.

This note describes the differences obtained in practice between I.Qs. derived from two older tests, based on mental ages, and standardised scores, derived from an NFER test, based on percentiles calculated for each month of age.

II.—Procedure.

Two mental age scales, the Simplex Junior Intelligence scale (Richardson) and the Essential Junior Intelligence Test Form A (Schonell and Adams) were used, together with the NFER Primary school Verbal Intelligence Test No. 1 (Pidgeon).

The Simplex, first published in 1932, has been standardised on the whole child population of a large (unspecified) town—some 12,000 children between the ages of 6 and 14 years. The Essential has been standardised on 3,200 pupils between the ages of 6 and 14 years. In both these tests mental age is taken as that chronological age for which the test performance is average (i.e., the regression of test performance on age is used).

The NFER test has been standardised on 6,419 children between the ages of 8·3 and 9·11, calculating percentiles for each month of age and converting into standard scores via the normal curve.

The three tests were given to three 'A' classes in three different Dudley schools in such an order as to spread practice effects equally between the tests. All tests were given at 9.30 a.m. with a few days' interval between tests. In all, 128 children between the ages of 9·6 and 10·5 years did all three tests.

III.—Results.

	Mean I.Q.	S.D.
Essential A	117	12
Simplex Junior	117	13
NFER Primary Verbal No. 1	108	11

These results confirm teachers' impressions that the mental age scales yield higher I.Qs. than standardised scores. This was particularly marked for children over I.Q. 130 :

NFER Standardised Scores	Average Discrepancy Essential I.Q.	Average Discrepancy Simplex I.Q.
130–139	+24	+20
90–129	+7	+8
80–89	0	−6
70–79	−7	−7

(Figures including an additional 29 ' B ' class children.)

IV.—Discussion.

These discrepancies might be accounted for by :

(1) The relatively low ceiling of the NFER test.

(2) Differences in standardisation groups used. As the three tests were standardised on different samples at different times the standardisation samples may differ in social class composition and in test skills, and test sophistication.

(3) Differences in the method of standardisation.

The NFER test is a good deal easier than the other two tests. Percentage scores at 10.0 corresponding to I.Qs. or standardised scores are :

	70	85	100	115	130
NFER	10	32	60	83	94
Schonell	10	32	53	69	81
Simplex	19	30	41	53	66

From this table it is clear that there are few items in the NFER test which discriminate above a standardised score of 130—on the other hand, no child did, in fact, reach the ceiling on this test. Schonell and Adams point out that their test is slightly unreliable at the extremes of the range of ability and that pupils scoring 90 per cent. should be given a more difficult test.

The second possibility—virtually a general increase in test performance since the years in which the older tests were standardised—seems a likely explanation of the middle-range discrepancy of some 7-8 points of I.Q. between the mental age scales and the NFER. The mean I.Q. of the Essential for Junior school children is given as 103, S.D. 16·1, which is very similar to the 1937 Binet mean and S.D. The mean I.Q. of the Simplex test is not given.

An increase in the average test performance of school children remains a possibility. The third possibility—differences in the method of standardisation—seems to be the most likely explanation of the discrepancies, however, particularly at the extremes. A standardised score of 130, for example, by definition is exceeded by no more than 2·3 per cent. of the child population, whereas a classical I.Q. has no such limitation.

It has also been demonstrated (Gulliksen, 1950, p. 287f.) that when the regression of score on age is used there will be a tendency for extreme scores to be exaggerated at the top and reduced at the bottom. This tendency is likely to be even more marked with less reliable tests than Simplex and Essential (split half reliabilities ·92 and ·95).

V.—Conclusion.

Two mental age scales have been shown to yield I.Qs. which deviate markedly from standardised scores, particularly above 130 and below 80.

VI.—Reference.

Gulliksen, H. (1950). *Theory of Mental Tests.* New York : Wiley.

CRITICAL NOTES

ON

Attainment Level on Leaving Certificate and Academic Performance at University by E. M. Gould and J. G. M'Comisky

This *Journal*, June, 1958

BY R. R. DALE

(*Lecturer in Education, University College of Swansea.*)

In this study the authors find the same moderate relationship between the university qualifying examination and degree results which has been found by Sutherland (1948) at Belfast, by Dale (1952, 1954a) at Swansea, by Sanders (1948a, 1953a) and Hohne (1951) in Australia, and even in U.S.A.—by Garrett. For ourselves and for the somewhat different conditions in Australia, this type of examination is as yet the best single predictor of success, though researchers would agree that prediction of pass or fail for those in the lowest third or perhaps quarter of the entrants is too hazardous to be satisfactory.

So far so good. The method is sound, and so is the finding that supplementary measures of prediction are needed. When, however, we look beyond the self-imposed limitations of this study at the many other factors which impinge on it, we see that, unless some of them receive mention, the findings will not appear in the right perspective.

The most important of these factors is the criterion. The authors use degree results ; but while they point out that achievement in after life might be a more valid criterion, if it were practicable to use it, they omit to give any consideration to the element of unrelability inherent in the degree results themselves. In doing this, they unwittingly project on to the Scottish Senior Leaving Certificate an implied criticism which really belongs to the universities. For, although many members of university staffs have a strange belief, that they can carry in their heads unchanged from year to year, and from one examination paper to another, an exact conception of a 40 per cent. pass mark, we know that the standard of marking is likely to fluctuate sufficiently widely to have an appreciable effect on the forecasting efficiency of the qualifying examination (Dale 1954b, Hohne 1955b, Hartog, and Rhodes 1935). Not only will it fluctuate according to the outlook of the individual examiner and the policy of separate departments, it will also vary according to changes in the average attainment of the student groups from year to year. As this variability of standard will affect the prognosticating power of any entrance tests which may be devised, it follows that one of our most urgent tasks is to attempt to improve university examining. One of the methods would be the use of high powered and refined objective tests as a small part of the examination in each subject ; another would be a further demonstration, by a committee of educational psychologists, of faults in present practice ; but further discussion of this would be out of place here.

A second point is that we are now at a stage in the research where the question, 'Prognostication—for what ?' needs to be considered carefully. There has been a perhaps justifiable tendency in early research to answer the question, as Gould and M'Comisky have done, by 'for university education,' but it is time we turned towards a more precise answer for groups of subjects or even single subjects. Special abilities play an important part in academic success at such a level, and if the necessary abilities for success in the university course are not adequately represented in the qualifying examination the craft may founder on the rocks ahead. As might be expected, a related group of subjects has generally higher predictive power for success in one of these subjects than have other subjects (Sanders 1948b). In England the student's choice of his Advanced subjects usually—but not always—automatically creates this weighting. In this connection more attention should be paid to clashes between the students' vocational interest and his abilities. Such cases, which are by no means negligible in number, depress the relationship between the qualifying examination as a whole and university performance. To cite only one example, a student gifted in languages deserted them at the university because of a vocational urge for social work ; she took an honours course in Social Science, where she had to take Statistics, though Mathematics had been her *bete noire* at school. She might easily have failed, and in fact, took a lower class degree than she might have done.

Sex differences in academic progress at the university are also worthy of inspection and it seems as if the authors of this study would have the material available. Evidence from a number of sources points to a general tendency for the best women to decline in comparison with the best men (Dale 1954c, Ministry 1950, Valentine 1932). If Gould and M'Comisky also found this tendency, fresh light would be thrown on the predictive effect of the qualifying examination. Other sex differences in academic progress might be found for the bottom third of the entrants and in individual subjects.

In conclusion, we might do well to ask ourselves where all this research work is taking us. It leads in the direction of perfection in prognostication, but this very concept has more than the usual element of illusion. Even if all the candidates for entry could be placed in accurate order of merit for university work, it would still be impossible to draw a line at any point (except an impracticably high one and that only for a few years), excluding all failures, because failures are psychologically essential to the present system of examining. As the standard is partly subjective, it will tend to relate itself to the average standard of the students ; the best will attain distinction and the weak will fail. If no students fail in a course for several years, the students will tend to slacken in effort and the examiners will also gradually re-adjust their conceptions of a 40 per cent. mark, until failures are again created. Herein lies another argument for the introduction of objective tests.

REFERENCES.

DALE, R. R. (1952). The prognostic value of the university entrance examination. *Brit. J. Educ. Psychol.*, 22, 124-139.

DALE, R. R. (1954). *From School to University*. London : Routledge and Kegan Paul, (*a*) pp. 31-105, (*b*) pp. 60-105, (*c*) pp. 36-38.

GARRETT, H. F. (1949). A review and interpretation of investigations of factors related to academic success in colleges of arts and science and teachers' colleges. *J. Exper. Educ.*, 18, 91-138.

HARTOG, P., and RHODES, E. C. (1935). *An Examination of Examinations*. London : Macmillan, pp. 55-56, 71.

HOHNE, H. H. (1951). *The Prediction of Academic Success*. Melbourne : Australian Council for Educational Research.

HOHNE, H. H. (1955). *Success and Failure in Scientific Faculties of the University of Melbourne*. Melbourne : Australian Council for Educational Research, (*a*) pp. 98 seq., 120 seq., (*b*) pp. 114 seq.

MINISTRY OF EDUCATION (1949). *Education in* 1949. London : H.M.S.O., p. 231.

SANDERS, C. (1948). *Student Selection and Acadmic Success*. Sydney : Commonwealth Office of Education, (*a*) pp. 48, 50, 68, 72, etc., (*b*) pp. 31, 33 seq.

SANDERS, C. (1953). University Selection : some psychological and educational factors. *Austral. J. Psychol.*, 5, 118-132.

SCOTTISH COUNCIL FOR RESEARCH IN EDUCATION (1937). *The Prognostic Value of University Entrance Examinations in Scotland*. University of London Press.

SUTHERLAND, M. B. (1948). *The Prediction of Success in the University Faculty of Arts by Means of Marks in the Northern Ireland Senior Certificate Examination*. M.S., The Queen's University of Belfast.

VALENTINE, C. W., and EMMETT, W. G., (1932). *The Reliability of Examinations*. University of London Press, p. 123.

STUDYING THE MEANING OF QUANTITATIVE TERMS

By D. SHEPPARD

(*Government Social Survey*)

In an interesting paper in this *Journal*, Cohen, Dearnley and Hansel (1958), discuss the implications of the vagueness of quantitative terms. One or two comments are perhaps worth making. In one table a very neat and systematic decline is shown in the means of values subjects give for "the number of chances out of a hundred that we will have a fine summer", that a weather forecaster has in mind when he says we are " certain to have a fine summer ", " likely to have a fine summer ", " probably will . . . ", " may . . ." and " it is most improbable that we will . . .". For this particular set of values, the writers omit to show how much overlap there is in the values given by the 125 students concerned, and do not make it clear if all the subjects even put the statements in the same (numerical) order. The fact that the statements were presented all at one time would increase the apparent systematic relationship, for the subject begins to have more information about what such statements are supposed to mean, as soon as more than one is presented. Though such an experiment is ' expensive ' in time and subjects, the way in which one can probably find most differences in interpretation is to give each subject only one statement to consider.

A further remark in their paper which deserves a comment is the suggestion that, whenever possible, teachers should indicate the range of values they have in mind when using a particular uncertain expression.

The difficulty here is that teachers are often dealing with matters which cannot be precisely quantified. Thus, if making a statement that a child is " very interested in his work ", an imprecise quantitative statement is being made about a quality which cannot be measured objectively. The same applies to graders and to all who are concerned with making assessments of quality. However, there are ways of determining what the person using such imprecise quantitative expressions means by them (Sheppard 1954, 1955a, Jones and Thurstone 1955), although the results would scarcely be communicable in the way that the meaning of statements studied by Cohen, Dearnley, and Hansel, might be. It is fair to add that the writer found teachers agreed no better than other subjects about the interpretation which should be put on such quantitative terms as ' very,' or ' rather,' even in cases where the teachers were working together in the same school.

Sometimes it is possible for teachers and graders to turn these judgments of qualities into marks, which can then be summed up. If those scoring systems used by judges of dairy products are in any way typical, however, it would appear that the logical difficulties involved in designing and applying this system may be such that as much is lost as is gained by such a procedure (Sheppard 1955b). Marks, however, may give a better impression to the uninitiated.

REFERENCES

Cohen, J., Dearnley, E. J., and Hansel, C. E. M. (1958). A quantitative study of meaning. *Brit. J. Educ. Psychol.*, 28, 141-148.

Jones, L. V., and Thurstone, L. L. (1955). The psychophysics of semantics. *J. Appl. Psychol.*, 39, 31-36.

Sheppard, D. (1954). The adequacy of everyday quantitative expressions as measurements of qualities. *Brit. J. Psychol.*, 45, 40-50.

Sheppard, D. (1955a). Descriptive terms and points systems for rating food qualities. *Food Res.*, 20, 114-117.

Sheppard, D. (1955b). Judgments of qualities with dairy products. *Dairy Ind.*, 20, 38-40.

SUMMARIES OF RESEARCHES REPORTED IN DEGREE THESES

The Value of Measures of Persistence in Selection for Secondary Education.

By CYRIL CRANE

(*Abstract of a thesis submitted for the degree of M.Ed., University of Manchester,* 1958)

THE PROBLEM.

The investigation arose from observations made by primary school teachers. They thought that if it were possible to take into account children's powers of 'sticking-to-a-task,' this would enable a more accurate prediction to be made as to the type of secondary education for which the children were most fitted.

Information was sought on the extent to which measures designed to test persistence with primary school children do measure a common factor, on the most effective methods of measuring persistence, and on whether persistence measures would significantly improve the predictive value of an existing battery of ability tests in selection for secondary education.

THE PILOT STUDY.

A preliminary investigation was carried out with 100 children aged 9·7 to 11·7 years to consider the first two parts of the problem. Scores for nine tests designed to measure persistence and three ratings on persistence were obtained, three of the tests using verbal material, three using numerical, and three using spatial material. These scores were correlated with each other and with scores from standardised tests in Intelligence, English and Arithmetic taken at 9 plus. A factor analysis was carried out using the Thurstone centroid (simple summation) method with rotation. A factor which could be identified as a persistence factor emerged and this accounted for 21·6 per cent. of the variance.

It was concluded that, with these children, the quality of persistence was not specific to particular situations but was generalised sufficiently to justify consideration of the rest of the problem.

THE MAIN INVESTIGATION.

From the twelve persistence measures used in the pilot study, three were chosen to investigate the value of persistence measures in selection for secondary education. The measures, which were chosen on the basis of high loadings on the persistence factor, low intercorrelations with other persistence measures, and simplicity of application, were: a word building test, a number square test, and a teacher's rating on persistence. Scores on these three persistence measures were obtained for 560 children in their last year in the primary school. Fourteen primary schools co-operated and it was felt that the children represented a reasonable cross-section of the community, unselected except for age. Marks obtained by these children in three tests of ability used in selection were made available.

CRITERION.

Two criteria of success were used:

(1) A criterion derived from orders of merit at the end of the first year in the grammar schools.

(2) A composite scale from ratings given by head teachers of both grammar and modern schools.

RESULTS.

(i) Inter-correlations of the persistence and ability scores and criterion (1) were obtained for a population of eighty-three children in two grammar schools and the multiple correlation coefficients calculated with the following result;

TABLE 1

		Multiple R
Int.+Eng.+P. Rating+Arithmetic+Word Building+Number Square	=	·857
Int.+Eng.+P. Rating+Arithmetic+Word Building	=	·854
Int.+Eng.+P. Rating+Arithmetic	=	·850
Int.+Eng.+P. Rating	=	·849
Int.+Eng.	=	·837

The increases contributed by the Number, Word and Arithmetic tests, respectively, were not significant. The increase resulting from the addition of the Persistence Rating was significant at the ·05 level.

(ii) When the two schools were considered separately, the Number, Word and Arithmetic tests again gave no significant increase to the multiple correlation in either case. With one school (population thirty-nine) the Persistence Rating also failed to give a significant increase, with with the other school (population forty-four) it was the best single predictor.

(iii) Inter-correlations of persistence and ability scores and criterion (2) were obtained for a population of 408 children in all secondary schools. The multiple correlations are shown in the following table :

TABLE 2

		Multiple R
Eng.+Int.+Arithmetic+P. Rating+Number Square+Word Building	=	·826
Eng.+Int.+Arithmetic+P. Rating	=	·826
Eng.+Int.+Arith.	=	·813

The addition of the Word and Number persistence tests gave no significant increase to the multiple correlation, but the addition of the Persistence Rating was highly significant.

Conclusions from the Main Investigation.

(i) The Persistence Rating would appear to predict success in the grammar school at least as well as does the Arithmetic test.

(ii) The persistence tests failed on all occasions to improve significantly prediction by the existing battery of selection tests.

The Effects of Different Instructions to Markers in the Marking of the English Essay.*

By PHILIP H. TAYLOR

(*Borough Road College, Middlesex*)

THE AIM OF THE INVESTIGATION.

The aim of the investigation was to discover the effects of different instructions to markers in the marking of the English essay for selection purposes. To this end twelve markers were asked to mark English essays using three different sets of marking instructions, called 'analytic,' 'Devon' and 'personal.'

PLAN OF THE INVESTIGATION.

Subjects: 1,072 children of selection age were asked to choose one of five titles and to write an essay for thirty minutes about it. Sixteen schools fully representative of a large urban area co-operated.

Markers: Twelve teachers, four men and eight women, acted as markers. Seven of them were graduates. Six of the markers had experience in the junior school; five in the grammar school, and the remaining one in the modern school only.

Marking Instructions: Three sets of instructions were drawn up and set against the same background of general impression marking. This back-ground was adopted from the marking scheme used by the Devon County Education Committee and reported by Wiseman.† The different instructions were designed to represent different approaches to the marking of the English essay, such that the 'analytic' instructions required attention to be given to the elements of grammar; the 'Devon' instructions to vitality and force in the writing, and the 'personal' instructions to the marker's own estimate of the requirements of the grammar school.

Sampling: From the 1,072 available essays a stratified random sample of 360 was taken such that there would be an equal number of each of the five titles and an equal representation of the sexes. From this sample, twelve groups of thirty scripts were selected and each marker was asked to mark one group using one of the instructions until he had marked three groups and used each of the instructions. After three months each marker re-marked a sample of the essays using with each sample the instructions he had used previously.

RESULTS.

The first marks were subjected to a four-factor analysis of variance and the re-mark provided by product moment correlation an estimate of the markers' self-consistency and of the reliability of the instructions.

TABLE 1

FOUR FACTOR ANALYSIS OF VARIANCE OF FIRST MARKING.

N=360

Source	S.S.	d.f.	M.Sq.	F	Sig.
Titles (T)	385·9	4	96·5	5·4	**
Markers (M)	683·3	11	62·1	3·5	**
Sex (S)	102·4	1	102·4	5·72	*
Instructions (I)	76·5	2	38·2	2·1	—
Residual	6107·4	341	17·9	—	—
Total	7355·4	359	—	—	—

---

* Summary of thesis in part fulfilment of the requirements for the degree of Master in Education at Manchester University, 1956.

† WISEMAN, S. (1949). The marking of English compositions in grammar school selection. *Brit. J. Educ. Psychol.*, 19, 200-209.

The results of the analysis of variance given in Table 1 show clearly that the instructions to markers have no significant effect on the marking. The differences which do exist are due either to chance errors or to interactions with other variables. That there should be significant marker differences is to be expected and to some extent desirable where the Devon scheme of using a team of markers is employed (Wiseman, 1949). Again, as expected, the girls were awarded generally higher marks than the boys. The difference between titles is probably a function of the ability to write an essay and to this ability being rewarded by the markers.

TABLE 2

SELF-CONSISTENCY OF MARKERS FOR EACH OF THE MARKING INSTRUCTIONS.

Markers	Instructions		
	Analytic	Personal	Devon
1	·89	·91	·77
2	·95	·93	·95
3	·85	·78	·98
4	·91	·88	·84
5	·79	·80	·97
6	·90	·88	·97
7	·96	·82	·93
8	·95	·93	·98
9	·73	·90	·87
10	·92	·89	·95
11	·87	·92	·93
12	·96	·94	·88
Mean	·89	·88	·92

Table 2 indicates that marker consistency was generally good, being over ·7 in every instance. Half the markers show no appreciable differences in consistency. However, three are relatively more consistent with the 'analytic' instructions (up to +·15), and three with the 'Devon' instructions (down to −·15). All the former markers are non-graduates, the latter graduates and each group prefers the method at which they are more consistent. The number involved is small, but the result is suggestive.

The results suggest that, against a marking background of general impression, radical differences in instructions to markers make no general difference to the marking or to the reliability of the markers. Further, the investigation indicates very tentatively that there may be some relationship between the instructions, the reliability of the marker, his preferred method of marking and his professional qualifications.

BOOK REVIEWS

AHMAVAARA, Y. (1957). *On the Unified Factor Theory of Mind.* Helsinki : Akateeminen Kirjakauppa, pp. 176, 750 mk.

The problem of universal factor-identification on an intuitive basis has claimed much attention in recent years. This book is concerned with the objective comparison of factors arising from similar research studies, and with the utlimate possibility of a unified factor theory.

The early sections offer a general description of the factor analysis theory based on the Thurstone vector approach. Six linear vector models are obtained, and each is shown to have sharp implications for psychological experiment. Any actual factor analysis is made as a finite approximation to one of these models. Against this background, the need for comparing factors arising from similar analyses is clearly shown, and the general theory of rotation and transformation analysis is introduced.

The later sections give a detailed description of the method of transformation analysis and its application to the comparison of factors present in some established studies, such as those of French, Adkins and Lyerly, Thurstone and others. The concepts of factor-invariance and factor-fission are introduced. Known factors are ranked in three certainty levels for their psychological soundness and consistency. It is of interest to note that number and mechanical dexterity factors appear to demonstrate the highest consistency, while verbal and reasoning factors tend to show greater divergence in different studies. Finally, an attempt is made to construct a factor pattern which reflects a unified theory of mind.

The application of neat mathematical transformations to psychological problems often necessitates consideration of information which has been omitted. To compare factors arising from different analyses by means of common tests would seem to ignore the contribution of all other tests in either battery. The factors under comparison have been mathematically generated by separate and differently-constituted test environments, and how much can they be reasonably expected to concur ? Is it possible that number or manual dexterity factors will concur more highly because, by their very nature, they are less susceptible to the influence of other tests ? Again, actual methods of factor analysis differ among themselves. Would not these, as well as rotations, contribute to differences in related factor patterns ? Perhaps, also, some form of rotation should be applied to improve the interpretation of the coefficients in the final comparison matrix ? Lastly, what level of invariance is essential to a sound and consistent factor? These and other problems are generated by the book. The subject is worthy of the most careful consideration.

D. M. LEE.

COHEN, J. (1958). *Humanistic Psychology.* London : George Allen and Unwin, pp. 206, 18s.

'Essays in Humanistic Psychology' would be a less inaccurate title than 'Humanistic Psychology' for a book which lacks the unity, comprehensiveness, and systematic treatment the latter title implies. Certainly, an attempt to deduce the nature of Humanistic Psychology from the contents of this book would yield little beyond certain dislikes—of nonsense syllables, rat psychology, factor analysis, for example, and a liking for the clinical approach, the data of personal experience, and the work of Lewin and Piaget, Sullivan and Freud. Hardly enough for another 'ism !

The actual text, if notes, references, and index are excluded, consists of less than 180 pages, with about 400 words to the page. Not too much for a book which, the author claims, 'might be equally described as "an elementary psychology for advanced students" or as "an advanced text for beginners"'; not too much, even if the whole book had been strictly to the point. Actually, the book is a collection of ten essays of uneven value that do not obviously depend on each other. The reader would be well advised to read those that interest him, without feeling obliged to read the rest.

The best essay is that entitled ' Psychological Time ' ; this could well have been longer at the expense of irrelevancies such as gibes at other ways of thinking than those the author prefers. Some of these, like the remarks on pp. 110-1, about psychological text-books as regurgitations of previously learnt nonsense syllables, are not only irrelevant but in poor taste. Others, like the ' intelligence test ' on page 8, simply puzzle. Ten questions are taken from the Book of Job, such as ' Canst thou bind the sweet influences of Pleiades, or loose the bands of Orion ?,' to be answered, ' Yes, No, or Don't Know .' One is left wondering whether this is just a joke or whether the author supposes that rhetorical questions are suitable for Intelligence Tests.

H. E. O. JAMES.

GABRIEL, J. (1957). *An Analysis of the Emotional Problems of the Teacher in the Classroom.* Melbourne : F. W. Cheshire, pp. xvi+224, 35s.

This book is an account of a thorough and painstaking empirical research on the factors causing strain in teaching, and on the emotional problems which beset the teacher in carrying out his duties.

There were two main phases in the enquiry. In the first, a sample of 162 teachers filled in an ' open ' questionnaire designed to elicit the major sources of frustation and annoyance occurring in teaching, as these spring from children's behaviour in the classroom, their attitude towards work, their personality traits, etc. The respondents were also asked to specify the occasions for feelings of depression and elation, and to make suggestions for changes in the school curriculum.

In the second part of the enquiry the replies to the first part were used to construct a second questionnaire which was answered by seven hundred and thirty-six teachers. The aim of this was to consider the effects of certain chosen variables on teachers' reactions, the relative frequency of strain or pleasure in relation to various situations and teachers' reactions to various reforms. By using the personal data supplied by the teachers, an analysis of the data was made according to types of school, sex of teacher, years of experience, size of classes taught, and type of district.

A mass of detailed results is reported and, while nothing very new comes to light—in general the factors causing strain or satisfication in teaching are well-known—the empirical documentation is valuable in providing irrefutable evidence of the adverse conditions under which a large proportion of teachers work, and of the need to help students in training for teaching to achieve a deeper understanding of their pupils and themselves. A valuable part of the book is the psychological and philosophical discussion of the problems with which the study deals. The approach is a balanced and sober one. Many of the conclusions are sobering too, in that it is clear that a majority of the profession are but little persuaded as yet of the desirability of many of the reforms of the educational system currently advocated by those who have given most thought to the matter. The need to lengthen teacher training and to cultivate insightful attitudes among students is stressed. The book is a valuable addition to the comparatively small number of empirical studies in this difficult but vital field.

BEN MORRIS.

GOINS, J. T. (1958). *Visual Perceptual Abiltiies and Early Reading Progress.* Supplementary Educational Monographs, No. 87. University of Chicago Press (Cambridge University Press), pp. x+109, 21s.

Though reading clearly necessitates the ability to perceive small details of shape correctly, no evidence has heretofore been obtained of any close correlation between reading achievement and scores on tests of discrimination of shapes other than those of words and letters. Moreover, there is considerable doubt as to whether general factors of perceptual ability such as those extracted by Thurstone have any relationship to reading ability. Neither has training in tachistoscopic perception been proved to produce much improvement in reading, except in the rather problematical investigations of Renshaw.

In this monograph, however, Dr. Goins demonstrates that small correlations were obtained between the reading achievement of first grade children and their scores on various perceptual tests. And there was a correlation of as much as 0·5 with a pattern-copying tests and one of reversal tendencies. The latter may, of course, be related to the appreciation of order and orientation which is certainly important in reading. But many of the tests (especially the pattern-copying test) showed appreciable correlations with intelligence test scores. Since no calculation was made of correlations between perceptual tests and reading achievement with intelligence partialled out, the degree of association remains doubtful. Different types of perceptual skill appeared to be involved in the different tests ; but two general factors were extracted, which the author called persistence and strength of closure, by comparison with those obtained by Thurstone. The second factor appeared to correlate with reading achievement also.

Dr. Goins also trained a group of first grade children in the tachistoscopic perception of digits and geometrical figures. But, although the children were able to perceive more digits after training than before, their reading achievement was unaffected.

This is a useful study as far as it goes, but it leaves many questions unanswered. Much more investigation is required as to the psychological nature of the perceptual skills involved in the tests, and the part they play, if any, in reading performance. There is the possibility that other factors may have influenced both test scores and also reading achievement ; or indeed, that superiority in the former may have resulted in part from superiority in the latter. There is some evidence for this supposition in that correlations increased between the first reading test, given after three months in school, and the second, given after nine months. M. D. VERNON.

INHELDER, BARBEL, and PIAGET, JEAN (1958). *The Growth of Logical Thinking.* (Translation by Anne Parsons and Stanley Milgram.) London : Routledge and Kegan Paul, pp. xxvi+356, 32s.

This work is concerned principally with the thinking of adolescents and the development of 'formal reasoning.' Piaget outlined this development in the *Judgment and Reasoning of the Child*, stressing the verbal character of reasoning at this level. Thus, the verbal problem, " Edith is fairer than Susan ; Edith is darker than Lily ; which of the three is the darkest ? " was instanced as requiring a more advanced type of thinking than the seriation of length since the elements of the series are presented in a purely verbal medium. However, the present experimental work, carried out by Inhelder, shows convincingly that formal reasoning is not confined to situations where the subjects work only with pencil and paper, but extends to all those occasions where they are required to establish some kind of structural order not merely on the raw experimental data but on the classifications and seriations to which they give rise. In this sense, formal reasoning is a higher order 'operation' whose elements are concrete 'operations' (classifications and seriations), the latter having been elaborated between the ages of 7 and 9 or 10.

The first fifteen chapters of the book are devoted to a series of experimental studies in which Inhelder applies the now familiar 'clinical method' to the study of how children proceed at various ages when required to educe the physical (or chemical) laws governing the operation of a variety of experimental apparatus. In a typical experiment, a selection of bars can be made to overhang a pool of water. They vary in length, in metallic composition, in cross-section and in thickness. An additional variable is the weight of the 'diver' who stands at the extreme end of the bar. Each of these five variables is a determinant of flexibility. The child is required to predict whether the bars will be flexible enough to reach the water level and to explain why. When asked to demonstrate the influence of weight, children at the level of 'concrete intelligence' will frequently vary length as well as weight to make doubly sure of proving their point, in effect proving nothing. Only at Stage III (11-15, i.e., the level of formal reasoning), is the problem solved systematically, by altering each variable in turn while holding all the others constant.

Piaget analyses the replies made by the various subjects in the concluding section of each chapter, and more fully, in chapters XVI and XVII. The stages involved in children's reasoning at various ages are translated into logical notation, bringing out the structural similarity of trains of argument involving different content.

To handle the method of ' all other things equal ' the adolescent must do more than read off experience as given. He must first formulate the possibilities in a systematic way, and then verify whether these possibilities are, in fact, realised.

Thus, Piaget's final definition of formal reasoning is ' the sub-ordination of reality to possibility ' or ' hypothetico-deductive ' reasoning.

Later, Piaget lays considerable stress on his exposition of that system of logical transformation which frequently governs the use of combinatorial systems. Non-logicians will certainly find this part of the work very hard going and logicians may well question Piaget's too easy transition from the qualitative (reciprocity) to the quantitative (proportion).

Criticisms which have been made of earlier works by Piaget apply even more forcibly to the present work. Although the number of children involved in the fourteen experimental situations is given as 1,500, no more than one or two examples are given to illustrate each level in each experiment. Frequently, the distinction between an example supposedly illustrating concrete intelligence and one illustrating the formal level is extremely subtle, or manifestly influenced by the instructions of the experimenter. Again, the functioning of formal reasoning may be less general than he suggests, i.e., the use of hypotheses or of arguments involving a considerable degree of complexity is dependent on the familiarity of the particular material and varies not only from individual to individual (at equivalent ages) but also from situation to situation within the same individual, and moreover, is considerably influenced by teaching. While Piaget does not deny that schooling plays a part, he insists that this must not be exaggerated. Yet recent work has shown conclusively that effective mental functioning is overwhelmingly influenced by educational experience.

For all that, this work is the first, and so far the only detailed experimental study of what is involved in rigorous conceptual thinking. Both the mode of approach (the ingenious experiments devised by Inhelder) and the main lines of interpretation (the logical exposition of the underlying thought framework given by Piaget) deserve the close attention of educationalists. They point the way to the formulation of psychological criteria which might govern the adaptation of content and method in education to the child's level of development.

The translators and publishers deserve our thanks for making this book available in English comparatively soon after its original publication. Although there is evidence of undue haste in the correction of proofs, and the translation is often over-literal, it is usually accurate. The provision of diagrams to illustrate the various experimental situations is a considerable aid to understanding. E. A. LUNZER.

LOVELL, K. (1958). *Educational Psychology and Children.* London : University of London Press, pp. 272, 18s.

The primary aim of this book is to provide a comprehensive and up-to-date text-book for teachers in their intiial training and an introductory course in educational psychology for those taking a higher qualification in education. It is certainly both up-to-date and comprehensive for, although short, it covers all the topics one expects to find in a book of this title. It deals with motivation, learning, child development, mental measurement, adjustment and maladjustment and so on. There is a final chapter on elementary statistics. It would probably be true to say that all relevant work in this country up to 1956, and a great deal from America, has been referred to either in the text or in foot-notes. Controversial matters are not avoided but the writer himself does not take any extreme views, nor is he afraid to point to problems for which psychology has, as yet, no clear answers or to areas where research has failed to yield the information sought. At all times the relevance of psychological findings to classroom practice is pointed out so that one may hope for considerable transfer of what is read here into action in the schools.

In spite of these good features, however, it seems to the reviewer that this book will be unlikely to find its way, as a main text-book, into many departments of education and, even less so will it be used by training college students. This is because it is far too difficult to read. It is not that the language or sentence construction is particularly difficult but rather that new concepts are introduced so rapidly and with so little explanation that the inexperienced reader will be unable to assimilate them. Sometimes new and unusual concepts are brought in without any explanation at all, at others such explanation as is given comes not in the text but in a footnote, as, for example, when ' gestalt ' is defined on p. 86, ' topology ' on p. 100, ' vital capacity ' on p. 207, and ' barrier ' on p. 224. In fact, the sheer number of footnotes does much to make reading tedious, for over and over again the text is modified by an afterthought or addition which could better have been incorporated in the body of the script.

To have covered so much in so few pages has inevitably led to excessive condensation and to some superficiality. This would, perhaps, not matter if one could be sure that the reader was going to follow up a good proportion of the references listed, but while one would hope that advanced students would do just this, and so find the book a valuable guide to study, very few students in their initial training period have the time, the experience or the facilities to do so. Hence, there is a real danger that they will acquire an ability to talk glibly of cyclothymes, rhesus factor and internalised parental figures without really knowing anything about them.

In his aim to produce a new text for teachers in training, Dr. Lovell has been unsuccessful; what he has done is to provide training college lecturers and more advanced students with a very complete and accurate book of reference to guide their further reading.

L. B. BIRCH.

MOSER, C. A. (1958). *Survey Methods in Social Investigation.* London: Heinemann, pp. xiii+352, 35s.

Courses in educational psychology and psychometrics generally give considerable attention to such research tools as mental tests and examinations, but tend to ignore survey techniques such as random sampling, the interview and the questionnaire, although educational research is often as much sociological as psychological. This volume, based on lectures given at the London School of Economics, provides an extremely clear and comprehensive account of all that goes into survey studies, including a discussion of their designs ; their evolution in Britain ; the importance of random selection of samples and the improvement of precision by stratification, etc. ; non-respondents and what to do about them ; different types of material that can be collected ; questionnaires ; interviewing and the effects of interviewer bias and error ; coding, analysis and interpretation of results. Even educationists who do not engaged in research themselves should read this in order to equip themselves for evaluating surveys published by others, e.g., Gallup Poll results, investigations like *Early Leaving*, and so on.

ROBERTSON, J. (1958). *Young Children in Hospital.* London: Tavistock Publications, pp. xiv+103, 4s. 6d.

Mr. Robertson's small book, summarising for the lay reader the gist of his memorandum to the Platt Committee on the welfare of children in hospital, appears at a moment when a certain measure of reform, welcome in itself, has high-lighted the large untouched area of the problem. In it he re-states the dangers for children under 4 or 5 years of age of hospitalization without their mothers, emphasizing the spurious nature of the ' settling-in ' process which leads many in charge to mistake a deeply depressed, acquiescent child for a happy one ; and argues very cogently the case for accommodating mothers who can arrange to stay in the ward with their children. The companion film *Going to Hopsital with Mother* is now available.

Mr. Robertson faces honestly the difficult implications of this solution for anxious mothers, conscientious nurses and busy doctors, and for the family left at home, and concludes that none of these difficulties is generally insuperable and that the mental health of the child must come first. He appraises the merits and drawbacks of the various alternatives, starting with unrestricted visiting, looks at the special problems of fever hospitals and long-stay institutions, and challenges some established hospital traditions that still do violence to children's feelings for no sound reason. A book recommended for all concerned with young children.

T. W. MOORE.

SANDELS, S. (1956). *Utvecklingspsykologiska Beteendestudier.* Uppsala: Appelbergs Boktryckeri AB, pp. 273, kr. 25.

Gesell, Piaget and others have stimulated much work in the field of Child Study. Interest has centred on detailed observation of normal behaviour, on the one hand, and of carefully structured test situations on the other. Dr. Sandels describes an allied study concerned with the development of certain quantity operations in children between the ages of 1½ and 8½ years. Deliberate experimental studies of initially structured situations involving ordering, dividing-distribution, and assembling in receptacles were carried out with 460 children scattered through the age-range, and detailed records were made of each child's behaviour and reactions. On the day following the quantitative tests, each child was also given the Elmgren Translation of the Terman-Merrill test. Thus, the study is concerned with groups of children of different ages, and not with the follow-up of one group of children throughout.

An analysis of the behaviour recordings was made in an attempt to establish the type and distribution of behaviour occurrences in each half-year group. In common with Gesell, Piaget and Bühler, the author distinguishes three levels of development, these changing at three years and seven years respectively. Behaviour variations were seen to be few at the younger ages, and increasingly more with the older children. The actual trend of the thought shown in the experimental situations seemed to involve the three stages of syncretism, analysis and final synthesis, following closely Hebb's account of the perceptive processes, and the work of Janet and Piaget on the operational thought-processes. Some subsidiary work is offered on the relation of the I.Q. scores to such factors as behaviour, age, family groups and social background. As the sample of children is in no way representative of the corresponding child population, such work would have to be treated with the utmost care, and the results could be given weight only against that particular background.

The book is written in Swedish with an English summary. It is difficult to obtain all information, and interested research workers would do well to approach the author directly. Such items as the relation between ordering and classification, the length of time spent with each child, the nature of the testee-child relationship and the question of uniformity of instructions are but sparesly described. Yet the detailed form and results of the work would certainly be of great interest to child study workers in this country.

D. M. LEE.

SODDY, K. (edit.) (1958). *The Problems Arising from the Teaching of Personality Development.* Sociol. Rev. Monogr. No. 1, pp. 138, 12s.

The papers in this volume were presented at a conference held at the University College of North Staffordshire. Of the nine contributors one, Professor E. J. Shoben, came from Teachers' College, Columbia University, while the rest are members of the staffs of British universities.

The main theme of all the papers is the need for greater personal and social maturity both in the general population and, more specifically, among those members of the community whose function is 'helping others.' This maturity can be reached only through an understanding of the springs of inner life, the teacher and social worker learning to assist the personality development of others through an understanding of their own personal development. Many of the speakers made it plain that they accepted the " standpoint and conceptual tools of psychoanalysis as a basis for teaching personality development."

Whether methods which are appropriate to the social worker dealing with individuals so maladjusted that they need special treatment to help them to fit into society are also appropriate to teachers dealing with large groups of children, most of whom are not obviously maladjusted, is open to question. At the same time no one can doubt the value in a teacher of a sympathetic approach to the children based upon insight into their needs at their present stage of development and an understanding of the road they must take to maturity. How much teachers and those responsible for training them will acquire from reading these papers is doubtful. If they inspire some to make a study of personality development with a view to applying the knowledge to the education of their pupils, then the book may succeed in its purpose.

It is probably inevitable that papers of this kind, designed to be delivered orally, should lose some of their inspiration and sparkle when divorced from the personal contacts and discussions of the conference and presented in book form. It is a pity that the proof-correcting has been badly done and that the large number of misprints tends to induce a feeling of annoyance in the reader. K. M. EVANS.

TEMPLIN, M. C. (1957). *Certain Language Skills in Children.* Minnesota University Press and Oxford University Press, pp. xviii+183, 32s.

This is a good piece of work with clearly defined aims, systematically pursued. The design and procedure of the investigation were planned with care and carried out with a clear awareness of the difficulties inherent in recording and analysing the speech of children and a recognition of the precautions necessary in interpreting the results.

Four-hundred-and-eighty children between three and eight years of age were studied, each age-group, consisting of equal members of boys and girls, being a representative sample of the socio-economic levels of U.S. urban population. Four aspects of language were measured : articulation of speech-sounds ; discrimination of heard speech-sounds ; length and complexity of spoken sentences ; extent of vocabulary recognized and in use.

The first result of interest is the extent to which the children had, at successive ages, mastered the forms of speech. By eight years of age, articulation and speech-discrimination were normally almost completely accurate, while three-year-olds attained mean scores of between 50 and 60 per cent. Again, at eight, the mean score for grammatical accuracy was 75 per cent., while at three it was already 48 per cent. These figures would no doubt surprise those who take a poor view of the speech of children to-day.

As between the sexes, Dr. Templin's results do not wholly bear out the common belief that the linguistic development of girls is more rapid than that of boys. Although in general the girls obtained higher scores, 'the differences are not consistent and only infrequently statistically significant.' Dr. Templin makes the interesting suggestion that during recent years, the differences in language ability may have become less marked because of the tendency ' toward a single standard in child care and training in the last few decades.'

In comparing socio-economic levels, Dr. Templin finds, as might have been expected from much other recent work, that differences in socio-economic status correspond to statistically significant differences in intelligence scores and in nearly all the language abilities here measured.

Most striking of all is the inference she makes from a comparison of her results with those of other workers twenty-five years ago. The general patterns of development, with age, in sentence length and complexity and in the use of the various parts of speech are confirmed, but there is an ' increased loquacity of children in child-adult situations and a tendency for children of the same age to use more mature language than they did twenty-five years ago . . . in that at any age fewer one-word remarks and simple and incomplete sentences are uttered and more complex sentences and adverb clauses are used . . . The increased talkativeness found in the present study would seem to reflect an increased amount of adult language in the child's environment, whether as a result of increased viewing of T.V., more inclusion in family activities, general permissiveness toward the child's behaviour, or other factors.' These are hypotheses that future workers can hardly ignore.

M. M. LEWIS.

PUBLICATIONS RECEIVED

The mention of a book in this list neither implies, nor precludes, a later review.

H. B. English and A. C. English (1958). *A Comprehensive Dictionary of Psychological and Psychoanalytical Terms.* New York and London: Longmans, Green, pp. xvi+594, 60s.

H. C. Dent (1958). *Secondary Modern Schools.* London: Routledge and Kegan Paul, pp. xiv+207, 15s.

F. Heider (1958). *The Psychology of Interpersonal Relations.* New York: John Wiley; London: Chapman and Hall, pp. ix+322, 50s.

P. Slade (1958). *An Introduction to Child Drama.* University of London Press, pp. xii+100, 8s. 6d.

I. R. Ewing and A. W. G. Ewing (1958). *New Opportunities for Deaf Children.* University of London Press, viii+149, 15s.

K. Soddy, *et al.* (1958). *The Problems Arising from the Teaching of Personality Development.* Sociological Review, Monograph No. 1, pp. 138, 12s.

A. Ferriere (1958). *Psychologic al Types and the Stages of Man's Development.* London: Heinemann, pp. x+165, 18s.

J. S. Macpherson (1958). *Eleven-Year-Olds Grow Up.* Scottish Council for Research in Education, Publ. No. XLII. University of London Press, pp. xxiii+184, 15s.

S. Goldberg (1958). *Introduction to Difference Equations.* New York: John Wiley; London: Chapman and Hall, pp. xii+260, 54s.

J. Maynard Smith (1958). *The Theory of Evolution.* Penguin Books, pp. 320, 3s. 6d.

W. Macneile Dixon (1958). *The Human Situation.* Penguin Books, pp. 443, 5s.

R. Winterbourn (1958). *Caring for Intellectually Handicapped Children.* Wellington: New Zealand Council of Educational Research; London: Oxford University Press, pp. viii+47, 4s. 6d.

H. L. Gordon (edit.). *The New Chemotherapy in Mental Illness.* New York: Philosophical Library, pp. xvii+762, $12.00.

M. D. Ainsworth and L. H. Ainsworth (1958). *Measuring Security in Personal Adjustment.* Univ. of Toronto Press, pp. xiii+98, 24s.

V. V. Kamat (1958). *Measuring Intelligence of Indian Children.* (3rd. edit.). Oxford University Press, pp. xvi+267, 17s. 6d.

P. Strevens (1958). *Aural Aids in Language Teaching.* London: Longmans, Green, pp. 44, 3s.

Katholieke Centrale Voor Studie en Research (1958). *Schoolgezondheidszorg* Tilburg: Uitgeverij Zwijsen, Pedagogische Publicaties No. 2, pp. 128, f.5.60.

B. Semeonoff and E. Trist (1958). *Diagnostic Performance Tests.* London: Tavistock Publications, pp. xv+176, 32s.

C. M. Fleming (1958). *Teaching: A Psychological Analysis.* London: Methuen pp. xiii+291, 28s.

F. J. Schonell, J. A. Richardson and T. S. McConnel (1958). *The Subnormal Child at Home.* London: Macmillan, pp. vi+167, 5s.

F. Van Steenberghen (edit.) (1958). *Psychology, Morality and Education.* London: Burns and Oates, pp. ix+128, 16s.

D. Richards (1958). *Offspring of the Vic.: A History of Morley College.* London: Routledge and Kegan Paul, pp. xi +316, 18s.

R. H. Ahrenfeldt (1958). *Psychiatry in the British Army in the Second World War.* London : Routledge and Kegan Paul, pp. xv+312, 35s.

H. S. N. McFarland (1958). *Psychology and Teaching.* London: Harrap, pp. 232, 10s. 6d.

P. Winch (1958). *The Idea of a Social Science and its Relation to Philosophy.* London: Routledge and Kegan Paul, pp. 143, 12s. 6d.

W. J. H. Sprott (1958). *Human Groups.* Penguin Books, pp, 219, 3s. 6d.

M. Fordham (1958). *The Objective Psyche.* London: Routledge and Kegan Paul, pp. vii+214, 25s.

P. Lafitte (1958). *Social Structure and Personality in the Factory.* London: Routledge and Kegan Paul, pp. xii+228, 24s.

L. P. Thorpe and A. M. Schmuller (1958). *Personality: An Interdisciplinary Approach.* New York: Van Mostrand, pp. v+368, 41s. 6d.

A. M. Clarke, A. D. B. Clarke, *et al.* (1958). *Mental Deficiency: The Changing Outlook.* London : Methuen, pp. vi+513, 45s.

VOLUME XXIX. JUNE, 1959 PART II

EXPERIMENTAL EXAMINATION OF SOME OF PIAGET'S SCHEMATA CONCERNING CHILDREN'S PERCEPTION AND THINKING, AND A DISCUSSION OF THEIR EDUCATIONAL SIGNIFICANCE*

BY E. A. PEEL

(University of Birmingham.)

SUMMARY. This paper describes four researches designed to test the substance of some of Piaget's schemata for the mental development of children. The first two investigations are concerned with haptic perception and spatial relationships in drawing, the third is devoted to the child's logical judgments, and the fourth to his moral judgments. Large numbers of children were put in Piaget's situations, the numbers ranging from thirty-two to sixty. The reliability of the judges' interpretation of the children's responses was measured in three of the inquiries by obtaining two independent sets of judgments and correlating the results. The validity of Piaget's schemata was assessed by comparing the development of the children's thinking and perception with the criteria of chronological age, mental age, and the total score giving each child's overall level of perception and thinking. This was done by adapting Guttman's technique to this problem. Some high reliabilities and validities were found and, in general, Piaget's conclusions were substantiated. A short discussion of the educational implications of Piaget's work concludes the article.

I.—INTRODUCTION.

DURING the last year or two several experiments on Piaget's lines have been carried out by research students in the Education Department at Birmingham University. In this paper four of these experiments are discussed. Two of them are concerned with children's perception and consist of repetitions of tests described by Piaget and Inhelder[8]. Mr. E. I. Page[6] repeated the experiment on Haptic Perception, (8, Chap. I), testing sixty children between the ages of 2:10 and 7 : 9. Miss E. Ferns repeated the experiment carried out by Piaget, Inhelder, Morf and Demetriades (8, Chap. II) on the child's treatment of elementary spatial relationships in drawing. She tested fifty-five children between the ages of 2 : 9 and 7 : 9 by asking them to copy the twenty-one shapes used by Piaget and classifying them according to his schema (8, pp. 55-57).

Mr. A. R. Lodwick set out to test Piaget's hypothesis of the three-stage development in children's logical judgments, namely by intuition, by concrete operations and lastly at the formal level. Some thirty children aged between 7 : 8 and 14 : 11 were given three short simple passages of historical interest—*Stonehenge*, *Alfred and the Cakes* and *Florence Nightingale*. After reading each passage they were asked three questions on its content, providing in all nine answers for each child which were then graded according to the stage of judgment they exemplified.

Mr. R. Loughran used a similar technique to investigate the same children's moral judgments by asking questions of a moral kind about brief passages on *Thermopylae*, *The Gunpowder Plot* and *The Battle of Cassino*. Here the three stages are described by Piaget as constrained, co-operative and autonomous levels of moral judgment.

---

* Based on paper given at Annual General Meeting of the British Psychological Society, Birmingham, April, 1958.

Mr. Page's material did not lend itself to quantitative analysis nor was it possible to arrange for other judgments than his own as to the level of the children's haptic perception, but in the case of the other three experiments, we were able to arrange for another judge (myself) to grade the drawings and verbal answers independently. Our consistency of judgment was then measured by correlating the pairs of sets of gradings.

Furthermore, by quantifying the levels of perception and judgment, logical and moral, we were enabled to say how far Piaget's three-decker schemata discriminated between different chronological ages. As we also had the mental ages available in the case of the work carried out by Messrs. Lodwick and Loughran, we were able to assess this discriminative validity against a criterion of mental age. Finally, by taking each child's total score in perception by adding the scores given to the twenty-one drawings, and the scores given to the nine answers to both the logical and moral questions, we could also use total level score as the criterion against which we could validate individual items.

By using Guttman's scalogram analysis[4] it was possible to express the discrimination of the Piaget schemata in the case of perception and logical and moral judgments, as a coefficient of reproducibility taking chronological age, mental age and total level as criteria.

I shall, therefore, concentrate on the last three experiments on pictorial space and logical and moral judgments. But, before doing so, a brief comment on the experiment on haptic perception and its educational significance will be helpful, particularly as this experiment logically precedes the study of pictorial space.

II.—HAPTIC PERCEPTION.

The principal result of Mr. Page's work, described fully elsewhere,[6, 7] was a clear confirmation of the general findings of Piaget and Inhelder. This is to the effect that topological shapes are recognised with greater ease than euclidean shapes by younger children.

Topological forms involve only the relationships of proximity, separation, order, enclosure and continuity, whereas euclidean shapes involve metric and directional relationships as well. Thus, for young children the triangle, diamond and circle forms are all equivalent to touch. We may put it in another way, that the geometrical abstractions implied in the straight line, in the angle and in dimensions are not appreciated by young children until the more primitive perceptual forms, possibly closely connected with the more undifferentiated movements and vision the young child is capable of exercising, have been experienced.

Educationally it seems clear from these experiments that experience of topological forms must precede any attempt to make a child aware of euclidean forms. Indeed, if these are given to the young child all he will perceive in them are their more general topological features. In the next section where I describe the work on children's 'Pictorial Space' we shall see that we are justified in extending this claim to visual as well as haptic perception.

A further educational corollary relates to the actual manipulative, drawing, reading and writing skills we might expect of young children. It seems that if we present too early material which implies an ability to perceive and execute euclidean forms, then we are certainly inviting failure. The earlier manual and visual skills expected of young children should be those that involve only the main topological ideas of proximity, separation, enclosure, order and continuity.

Lastly, by giving more experience to children of such movements and perceptions we may be preparing them more fittingly for later classroom experiences of the usual geometrical kind. But the point as to whether the development from the earlier topological to the later geometrical stage is largely one of maturation of muscular and sensory control, mainly independent of exercise, or whether it can be accelerated by providing plenty of experience is still undecided.

III.—SPATIAL RELATIONSHIPS IN DRAWING.

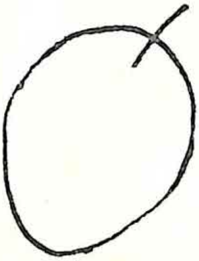
The twenty-one simple shapes used by Piaget (8, p. 54) in the drawing test are chosen to represent geometrical features of *contact*, *intersection*, *inside*, *outside*, *straight* and *curved lines*, *direction* and *orientation*, and they include closed figures, circles, triangles, squares, rhombuses, the ellipse, and crosses.

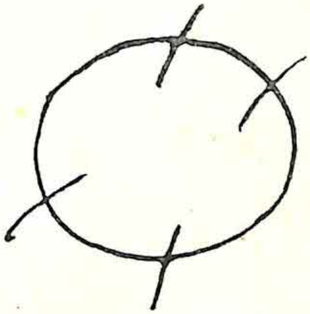
In the child's development of capacity to copy such simple shapes, Piaget suggests the following important stages :

Stage O (2 : 6—2 : 11 years)—scribble.

Stage Ia (3 : 6—3 : 10)—still of a scribble nature but open shapes are distinguished from closed figures (closure is probably the first feature recognised after apprehension that there is something to be copied).

Stage Ib (3 : 6—4 : 0 years)—the beginnings of real drawing but only topological relations are depicted. The triangle, circle and square are all given as closed continuous figures—occasionally there is the slightest symbolism such as :





for a triangle and square respectively.

Stage Ib—Ia—curved lines begin to be distinguished from straight lines. The rectangle is frequently drawn more correctly than the square.

Stage IIa—shapes begin to be distinguished by angles and dimensions.

Stage IIb—more complex circumscribed, tangential and intersecting figures involving notions of contact, intersection, inside and outside are correctly depicted, but not always metrically correct.

Stage III—when all figures are correctly drawn, including their relative size and dimensional properties.

We found it helpful in assessing the fifty-five sets of drawings produced by the pupils in Miss Ferns' research to score intermediate phases. Piaget gives seven stages but we used an 11-point scale, in order to quantify the scoring, as follows :

Stage in Piaget's terms	0	—Ia	Ia	—Ib	Ib	—IIa	IIa	—IIb	IIb	—III	III
Quantification	0	1	2	3	4	5	6	7	8	9	10

Miss Ferns first graded the drawings and then I graded them quite independently. On the above numerical scale I tended on the whole to be 0·5 units more severe, but the relative agreement between us was very high indeed. The average coefficient for twenty-one shapes (n=55) was ·89 and the lowest correlation was ·83 with a standard error of ·042.

This is important since Piaget himself gives no evidence of the extent to which other judges would interpret the material produced by the subjects of his investigations, and the lack of evidence has promoted some criticism of the value of Piaget's conclusion.

The next problem concerns the discrimination of Piaget's scheme of analysis, as set out on page 91, with respect to the criteria of chronological age and mental age and finally the validity of the twenty-one items against a criterion of total score. As we shall see, this latter problem is really a special case of discrimination.

The first step was to set out the 53* × 21 gradings on the 11-point scale, 0–10, in a rectangular tabular form suitable for inspection of the data along the lines of Guttman's scalogram analysis.[4] This was done by arranging the fifty-three sets of twenty-one scores vertically in ascending order according to the ages of the children. The grades chosen were the averages between Miss Ferns and myself.

Inspection of this table showed that, although there was a clear cut progression of scores with age, eleven grades were too fine in that they permitted far too much intermingling of grades according to age to allow of a reasonable coefficient of reproducibility.[4]

For the non-technical reader we may illustrate what is involved by an artificial example. Supposing we have fifteen people, A–O, arranged in order of age from left to right and that each scored 0–10 for his drawings as shown in the second line. Clearly, there is a tendency for high scores to be associated with older people, but the system of grading on the 0–10 basis may be too fine since there are are reversals (e.g., C scores 3 and D scores 1).

Example :

TABLE 1

Youngest	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	Oldest
Drawing Score	1	0	3	1	3	1	2	6	4	5	5	7	10	9	8	11 categories
Combined Categories	1	1	2	1	2	1	1	3	2	3	3	3	4	4	4	4 categories
	1	1	1	1	1	1	1	2	2	2	2	3	3	3	3	3 categories

We are attempting to find a system that admits of as few reversals as possible, since our whole problem is to say whether Piaget's scheme reveals real grades by age. We need to combine the categories until a more convincing rise with age is shown, involving only a few reversals of order. Guttman's essential technique in attitude questionnaires is to count these reversals. If there are none, then the attitude test is said to give results which are reproducing perfectly the information required. The analogy with the Piaget scheme of assessment of the drawings can perhaps now be better appreciated.

---

* Two children did not complete all the drawings and had to be dropped from that part of the analysis.

Suppose in order to achieve this we combine the eleven categories in the hypothetical example as follows:

0, 1, 2 combined and called 1
3, 4 combined and called 2
5, 6, 7 combined and called 3
8, 9, 10 combined and called 4

This means that we have decided that our judges cannot, in fact distinguish consistently between what they originally thought were the different categories of 0, 1 and 2, and similarly for the other groupings 3, 4, etc. The newly combined categories are given in the third row of the table. Things now look a lot more convincing, for there are only three reversals which all concern the new category 2, two of which are mixed with category 1 and one with category 3. We have, in fact, three wrong placings out of fifteen. This is not much and means that the evidence would support Piaget's schema on page 91. to the extent that there are only four stages in drawing maturity that progressively operate over the age range 2 : 9 to 7 : 9. We could make reproducibility perfect by combining as follows:

0, 1, 2, 3 combined and called 1
4, 5, 6 combined and called 2
7, 8, 9, 10 combined and called 3

This is done in the fourth row of the table and would represent the grading scheme in three grades to be insisted upon by a person setting a very severe test to Piaget's claims.

The example is hypothetical but let us now turn to the actual material obtained in this research. There are, of course, twenty-one sets of scores instead of one and fifty-three persons in each instead of the fifteen of the hypothetical example.

It was decided to combine the eleven categories as follows:

0, 1, 2, 3, 4 to make new category 1
5, 6 to make new category 2
7, 8 to make new category 3
9, 10 to make new category 4

This, in fact, allows for the original Piaget categories of I (a and b combined, also with 0*), IIa, IIb and III, which, he suggests, can account for the major changes in children's drawings as they grow older. When the grades were so combined, the following number of misplacements of gradings according to age were obtained:

TABLE 2

Drawing	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	Total
†Cat. 1	3	4	4	5	4	4	2	6	5	3	3	3	4	4	5	3	4	5	5	5	4	
†Cat. 2	3	4	4	6	6	8	3	8	4	4	4	3	5	5	7	5	5	8	8	8	4	
†Cat. 3	2	4	3	3	3	6	5	4	3	3	2	2	4	2	3	3	3	6	5	6	3	
†Cat. 4	2	3	3	2	1	2	3	1	2	1	1	2	3	1	1	1	2	3	2	3	3	
Total	10	15	14	16	14	20	13	19	14	11	10	10	16	12	16	12	14	22	20	22	14	315

* There were, in fact, only one or two drawings in our population of 1,113 drawings which receivd a zero grade both by Miss Ferns and myself.

†—Misplacements out of 53 places for each item.

Percentage errors, according to a criterion of age, in the whole testing situation of twenty-one items copied each by fifty-three children :

$$=\frac{315\times 100}{53\times 21}=28{\cdot}3\%$$

This would yield a Guttman coefficient of reproducibility of 71·7 per cent., whereas Guttman in scaling attitude questionnaires, insisted upon a minimum of 85 per cent. Guttman's condition is very stringent and not often met, and in this work we might say that 75 per cent. is a reasonable figure to take. If so, it is clear that we support Piaget's contention that children's drawings do progress according to the age of the children in the sequence he suggests. We may also comment upon the age ranges for each phase. When we examine our full table of 21×53 entries in the four category gradings, we find that the rough zones of demarcation between the categories are as follows :

TABLE 3

AGES ASSOCIATED WITH THE PIAGET PHASES.

Piaget phase	n	Median Age	Range of Age
1. (0 Ia, Ib)	13	4 : 2	2 : 9* —4 : 11
2. (IIa)	22	4 : 11	3 : 9 —5 : 4
3. (IIb)	10	6 : 6	5 : 7 —7 : 9
4. (III)	8	7 : 3	6 : 7 —7 : 7

*(No younger child studied)

Although the coefficient of reproducibility gives a fair case for Piaget's contention, we may ask whether there is any other order of the fifty-three children, that is, other than by age, that will give even greater reproducibility. Such an order is given by each child's *total score*, on the four-point grading, in all his twenty-one drawings.

When this is done, the 21×53 table of entries contains the same rows as before, but they have been reshuffled, up and down. The number of misplacements was counted as before with the following results :

TABLE 4

Drawing	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	Total
Misplacements.																						
Cat. 1	2	2	2	2	1	2	2	1	2	2	2	1	2	1	1	2	2	3	2	1	4	
Cat. 2	2	3	2	3	3	5	4	3	3	3	3	1	3	1	2	5	3	5	4	4	4	
Cat. 3	1	3	2	2	2	4	3	3	1	3	3	2	2	1	2	4	2	3	3	5	2	
Cat. 4	1	2	2	1	0	1	1	1	0	2	2	2	1	1	1	1	1	1	1	2	2	
Total	6	10	8	8	6	12	10	8	6	10	10	6	8	4	6	12	8	12	10	12	12	184

Percentage errors in the whole testing situation of twenty-one items copied each by fifty-three children.

$$=\frac{184\times100}{53\times21}=16\cdot5\%$$

This yields a Guttman coefficient of reproducibility of 83·5 per cent. and represents a very high confirmation that Piaget's sequence for the drawings of *individual* items correlates with the total *overall* drawing level of each chld.

What does this mean ? In one sense, it is a validation of the group of twenty one figures chosen to exemplify Piaget's thesis. But this is not all, for in another sense it can be taken, when compared with the association with age, to mean that Piaget's *sequence* of phases is fully established whereas his *age* placements are not so clearly supported. This, I think, is very important, since many workers are 'put off' by the age limits that Piaget lays down and in so doing come to criticise the sequence. This is a pity, for I do not feel that Piaget is fundamentally and rigorously concerned with the age limits, although he uses ages as an illustration to give the reader some idea of when the particular changes might take place. The really important thing is the *sequence* which remains constant although it takes place over different age spans.

This age difference is manifest in Piaget's writings and is obvious when as adults we tackle very difficult problems. We pass through a constant sequence but, according to whether the task is easy or difficult, be it in the perceptual or conceptual field, we find the sequence operating at *different* age levels, lower if the task is easy and higher if it is difficult.

IV.—LOGICAL JUDGMENTS.

The basis of Mr. Lodwick's method of inquiry[5] into the sequence of phases making up the development of the child's logical judgments is that given in Piaget's *Moral Judgment of the Child*. The three stories from history, together with the three questions asked after the reading of each passage, are given in Appendix I. The answers were analysed and scored on a five-point scale as follows :

1 for an intuitive answer.
2 for an answer between the intuitive and concrete stages.
3 for a concrete answer.
4 for an answer between the conrete and formal stages.
5 for an answer that showed evidence of fully formal and operational thought.

By intuitive thought Piaget had in mind thought not yet freed from perception and egocentricity. One aspect of the situation is focussed upon and others ignored. The child grasps only a single relation at a time, and this refers to concrete material ; hence his combined statements often contain inconsistencies. Here are instances from Mr. Lodwick's research (the questioner's remarks at the time are given in parenthesis).

Question : Could Alfred cook ?
A. Answer : Yes (Why ?) Because he's the King.
(Can every King cook ?). No.
(Why could Alfred cook ?) Because he can fight.

This answer scored 1 on the scale.

B. Answer : No (Why not ?). He forgot all about the cakes. He was a man and that's why he couldn't cook as well as a woman does. A woman is a proper cook.

This answer scored 2 since, although it is primarly at the intuitive level, it shows signs of power to co-ordinate relationships.

At the concrete level there is the capacity to combine simple relations between concrete objects, but not with abstract or formal problems. The child is not yet freed from centring and statements are judged in terms of their data rather than their formal relationship.

Here are two answers at the concrete level to the same question as before :

> C. Answer : I shouldn't think so, at least not very well. He didn't pay attention to the cakes, if he had been a cook, he might have known they'd been done.

The problem has been judged in terms of the data—burning the cakes. Therefore, Alfred could not cook. The relationship that Alfred was thinking about the Danes has been excluded. This answer received 3 points.

> D. Answer : I don't think so, because if he could have cooked, he'd have known how long it takes them to bake. (Are you sure he could not cook ?) I've never heard that he could cook. Kings are supposed to have servants for that. You couldn't be sure. I wasn't born in his time.

The answer was given 4 points because, although he is still under the dominance of the data (his first sentence), he can hold several relationships and he does go out of the data to his own experience.

At the stage of formal operations, the thinker is able to judge the statements formally and logically and is not put off by their content. Thought is said to be reversible, resulting in consistent argument. Hypotheseses are examined systematically. Here are two examples from the answers to the same question :

> E. Answer : No I don't think so, because in the olden days the King didn't have to know how to cook at all. It's like to-day, he has servants to wait upon him. If he could cook, he would know approximately when the cakes were done. He was thinking about the Danes but some instinct would have told him, if he could cook, when the cakes were done. We can't really tell.

> F. Answer : I do not know because if anyone could cook and had something on his mind he might still forget the cakes.

The analysis of the data followed similar lines to that described in the previous section. In all thirty-two pupils whose ages ranged from 7 : 7 to 15 : 0 took the tests and each answered nine questions, three on each of the three stories. The children were also tested for intelligence by the Raven's Progressive Matrices ; hence in this case we can use mental age as a criterion as well as chronological age and total score in the nine answers.

Two of us scored the answers independently and so we were able to say how consistently two judges could arrive at similar conclusions about the level of thinking revealed in the answers. The correlation coefficients between the judges' assessments of all the answers given to each story were as follows (n=96, i.e., 32 children, three answers each) :

TABLE 5

Story....................	r
Stonehenge	·67
Alfred	·78
Florence Nightingale	·62

The correlations are not as high as those obtained when the judges were grading perceptual material but are high enough to dispel the doubts of those who would reject Piaget's technique and results as following largely the subjective experience and feelings of one person. I am sure that with further more careful study on my part of the answers in terms of the Piaget scheme, the correlations would be inproved. Only in one or two cases did I find myself differing from Mr.Lodwick by more than two divisions on the five-point scale.

The next question was to settle how far the Piaget scheme differentiated between the children with respect to chronological age, mental age and total score.

Inspection of the table of the scores allotted on the five-point scale to the answers of the thirty-two children to the nine questions suggested that the categories two and four, giving borderline assessments, could well be combined with other categories. Three categories were now used corresponding to the original stages of intuitive, concrete and formal judgments suggested by Piaget. The frequencies of misplacements in the different categories, according to the three criteria used, are given in the following table :

TABLE 6

Number of Misplacements of Answers in the Three Categories Chosen.

	Stonehenge			Alfred			Florence Nightingale			Total Error	Percentage Errors	Co-efficient of Reproducibility
Criterion	1	2	3	1	2	3	1	2	3			
Chron. Age :												
Cat. 1	5	2	3	2	4	2	5	2	3			
Cat. 2	8	6	7	7	6	2	8	5	5			
Cat. 3	3	4	4	5	7	0	4	4	3			
	16	12	14	14	17	4	17	11	11	116	40·3	59·7
Mental Age :												
Cat. 1	4	2	3	2	4	2	4	2	2			
Cat. 2	7	6	7	5	5	2	6	5	3			
Cat. 3	3	5	6	4	4	0	4	4	2			
	14	13	16	11	13	4	14	11	7	103	35·8	64·2
Total Score :												
Cat. 1	2	2	2	3	4	1	4	1	2			
Cat. 2	4	7	4	7	7	1	6	3	4			
Cat. 3	2	6	2	4	4	0	2	2	2			
	8	14	8	14	15	2	12	6	8	87	30·2	69·8

The amount of intermingling of individual answers is much more marked than in the case of children's drawings and, even where it is least in the case of a coefficient of reproducibility of 69·8 per cent. against the criterion of total score, it falls far short of the high limits set by Guttman. However, there is evidence of such a trend of development as suggested by Piaget embodying intuitive, concrete and formal phases.

As in the earlier case the development is least marked by chronological age (C.Rep. 59·7), more marked by mental age (C. Rep. 64·2) and most marked by total score (C. Rep. 69·8). These results again emphasise that the important part of Piaget's writings on this subject concerns the *sequence* of the three phases rather than their firm attachment to particular age ranges.

Mr. Lodwick draws attention to the spread of logical maturity revealed in the nine answers given by each child. A child does not maintain a constant level of reasoning. This may be confirmed by reading across each row in the full table of three point scores[5]. He further demonstrates, by taking the *highest* of each child's nine scores and examining them for misplacements, that, in fact, a much more strongly marked and consistent development of reasoning can be revealed, as in Table VII.

TABLE 7

MISPLACEMENTS IN DIFFERENT CATEGORIES WHEN EACH OF THE THIRTY-TWO CHILDREN'S HIGHEST SCORE IS TAKEN AS HIS LEVEL OF REASONING POWER.

Criterion	Errors	% Error out of 32	Coefficient of Reproducibility
Chronological Age	8	25	75
Mental Age	6	19	81

This appears to be a valid measure, since the form of the different passages, the form of the questions and their relation to the passages, all vary and may limit some replies to certain parts of the range of reasoning. This last possibility is very marked in relation to the third question on Alfred. Also different passages and questions may motivate a child to varying degrees.

Overall, this research seems to provide definite support for Piaget's schema of children's logical development and suggests that with more careful selection of passages and questions, an even higher degree of confirmation could be reached.

V.—MORAL JUDGMENTS.

Piaget[9] suggests three levels of moral judgment in the child's response to situations involving his sentiments. At the lowest level there is judgment by constraint or authority where the child defers to outside authority. Then there is the level of co-operation and equality where he judges on a basis of mutual respect between persons and recognition of their equality. Lastly, there is the phase where there is autonomy of judgment and the child judges the situation by the general principle of equity. This last represents mature 'moral' judgment.

Mr. Loughran used three short passages of historical interest, all of which aroused moral issues; each was followed by three questions which further brought these issues to a point for the children tested. The stories were :

The Story of Thermopylae.
The Gunpowder Plot.
The Battle of Cassino.

They are given in Appendix II in full with the sets of questions associated with them. The same thirty-two children used by Mr. Lodwick were also used by Mr. Loughran for the research on moral judgments.

Instances of the different levels of moral judgment are provided in the following answers to a question based on "The Gunpowder Plot."

> Question : Was it right or wrong for King James I not to belong to his mother's church ?
>
> Answer (i) (at the level of authority) : He should belong to her church because his mother looks after him.
>
> Answer (ii) (at the level of equality) : It was right because he can pick his own church. Because he has a mind of his own.
>
> Answer (iii) (at the transition level of equality-equity) : Well, I think everyone should have his own opinion about religion and James was a grown man and free to choose. As he was King, however, I think he should have kept to his mother's religion to save upsetting the country.

The agreement between two independent assessments of the answers, expressed as a correlation coefficient, was ·74 for those on Thermopylae. A figure of ·81 was obtained in the case of Cassino, but here the agreement might be influenced by the fact that I saw the gradings given by Mr. Loughran. For the same reason no figure is given for answers to the Gunpowder Plot.

We may now discuss the discriminating power of the Piaget three-part schema against the criteria of chronological age, mental age and total performance (in the nine answers). When the answers are marked according to three categories corresponding to the levels of authority, equality and equity, it is clear from the following table that there are many more misplacements than in the case of the logical judgments.

TABLE 8

PERCENTAGE OF ERRORS (MISPLACEMENTS IN THE CATEGORIES) AND THE COEFFICIENTS OF REPRODUCIBILITY WHEN THE ANSWERS ARE MARKED IN THREE CATEGORIES.

Criterion	Percentage or Errors	Coefficient of Reproducibility
Chronological Age	51·4	48·6
Mental Age	46·9	53·1
Total Score	31·9	68·1

The figures obtained when chronological age is used as a criterion are, of course, quite unacceptable and, although the others exceed 50 per cent. reproducibility and also reveal the same general improvement as the criterion is changed from chronological age to total score, all coefficients suggest that the three-level schema suggested by Piaget did not discriminate in this test situation. However, we might find better discrimination by combining categories so that only two categories are used for each answer. Which pair are used—*authority-equality*, or *authority-equity* depends on individual items.

When this is done, the percentage of errors of placement in the two categories for each question, against total score as criterion, falls to 18·1 per cent., thus giving a coefficient of reproducibility of 81·9 per cent. This is high and on the basis of such a figure, we can assume that a real distinction can be made between the answers on these two level passes. In all cases it would seem that the real distinction is between answers made at the level of deference to authority and constraint and those at a level where equality between different persons is recognised, there being some indication in the more advanced answers that this idea of equality has extended to that of equity.

In making these observations, we should be clear that this limitation to two categories might, indeed, be a consequence of the particular reading passages and questions chosen for the test.

VI.—DISCUSSION OF THE EDUCATIONAL IMPLICATIONS.

It is helpful in the beginning to make two general points.

In the first place, Piaget's experiments and our confirmation of his results, in the particular experiments we choose, do not fully answer the question whether the development in children's perception and judgment proceeds as a result of pure maturation and readiness or is affected by experience and learning. Both factors probably enter. Other work by Churchill, [1, 2, 3] and unpublished tentative experiments at Birmingham suggest that experience is an important factor in promoting this development.

Secondly, it would appear that each phase of development is necessary for the emergence of later more mature phases.

If these assumptions are correct, it would seem that the more experience we can provide of materials, and reading and questioning at the appropriate level, the better.

With regard to the four researches discussed, the two investigations into the development of haptic and visual perception have clear-cut educational implications.

Young children are capable only of seeing, depicting and executing topological features and patterns. Awareness of and capacity to execute linear metrical and directional relations and operations follows at a later period. Very roughly the critical age range over which the transition is made from topological to euclidean awareness and capacities corresponds with the ages of entry into the infant school. In the case of slower developers the transition may occur much later. We may also at least hazard a guess that in the case of finer operations such as required in hand-writing, it may be later for all children than in the case of the more massive perceptions and manual operations required in drawing and handwork. Hence Piaget's results are relevant to the teaching of manipulative, drawing, writing and reading skills in young children. We must give plenty of experience that demands no more than topological awareness and capacities.

The study of children's logical judgments confirms the existence of the three-phase sequence suggested by Piaget. Educationally, it draws attention to the need for careful grading of passages for study by primary and secondary school pupils and sensitive choice of the scope of the questions to be used in connection with this study. The investigation described revealed how some questions in connection with the passage chosen were not capable of evoking the full range of response, intuitive to logical, As a tool, therefore, for discovering how far the child is progressing in his powers to deal with verbal situations, such questions and passages are defective.

It would also seem desirable to provide more opportunity for active answering on passages as opposed to passive reading. This would give the experience at each level necessary for progression to the next.

Lastly, although the numbers chosen (thirty-two) are too limited to make a firm comment on the ages of transition between different phases for grasping simple story material as chosen, it seems that for normal children the critical period is in the first three years of primary school life.

The findings in connection with the development of child's moral judgments, whilst not so conclusive, are at least suggestive of a sequence of help to teachers of religious knowledge, history and English.

First, if the purpose of an exercise is to discover maturity of judgment, material and questions need to be chosen so that they reveal the full range of phases. Secondly, younger children need plenty of experience of active reading and questions so that the latest and most mature phase of autonomous moral judgment may be more naturally and permanently achieved.

VII.—APPENDIX I.

STONEHENGE.

Stonehenge is in the South of England, on the flat plain of Salisbury. There is a ring of very big stones which the picture shows.* Some of the stones have fallen down and some have disappeared from the place. The people who lived in England in those days we call Bronze Age Men. Long before there were any towns, Stonehenge was a temple for worship and sacrifice.

Some of the stones were brought from the nearby hills, but others which we call Blue Stones, we think came from the mountains of Wales.

Question 1 : Do you think Stonehenge might have been a fort and not a temple ? Why do you think that ?

Question 2 : If the stones were rolled into position, what do you think has become of the rollers?

Question 3 : What do you think has happened to the missing stones ?

ALFRED AND THE CAKES (A Picture History for Boys and Girls by Richard Wilson).

King Alfred was an English King who lived long ago. In his time The Danes came from across the sea and tried to take the land from the English. King Alfred fought against them ; and, at first, they beat him, and he had to change his clothes and hide himself.

One story says that he came one day to a hut where a woman was baking cakes on the hearth. She let him sit by the fire to warm himself, but asked him to see that the cakes did not burn while she was busy.

He said he would, but began to think about the Danes and so forgot all about the cakes. When the woman came back the cakes were burnt. So she scolded him, not knowing that he was the King.

At another time, it is said, the King went as a harper into the camp of the Danes. He sang to them, and found out what they were going to do.

In time he was able to fight the Danes again, and this time he beat them.

Question 1 : Could Alfred cook ?

Question 2 : Did the Danes kill or capture every one of the English people ?

Question 3 : Did Alfred look like a King ? How do you know ?

FLORENCE NIGHTINGALE (A Picture Book for Boys and Girls by Richard Wilson).

About a hundred years ago there was a short war. The fighting took place in a land far away over the sea. It was Winter time, and our soldiers suffered very much from the cold as well as from wounds.

At that time there were no women nurses in the war hospitals, and very little was done to help the men who were sick and hurt.

Now a certain young English lady named Florence Nightingale had got herself trained as a nurse, because she had a great wish to help those who were sick.

So she was asked to go to the war and help the men who were sick and hurt. She said she would go at once, and she took other nurses with here.

They found that the men were not well looked after and that the hospitals were very dirty. But before long Miss Nightingale and her nurses made great changes.

The men were kept warm and were well fed. The nurses were very kind to them, and cheered them up so that they got better much quicker.

---

* A picture of Stonehenge is shown to the child.

The poor sick men loved Miss Nightingale very much. She used to go round at night with a small lamp in her hand to see if they were comfortable. So she came to be known as "The Lady with the Lamp."

Question 1: Why did they not put the wounded men in hospital in our own country?

Question 2: Why were the English war hospitals dirty? Why was that?

Question 3: Why do you think there were no women nurses in the army in those days?

VIII.—Appendix II.

The Story of Thermopylae (The World's Story by Elizabeth O'Neill).

The great King of Persia, Xerxes, made up his mind to conquer the stubborn Greeks and gathered a great army and invaded their country.

The Greeks knew that the Persian army must come through the mountain passes. They decided to make a stand at the Pass of Thermopylae, for if the Persians could get through that, there would be nothing to stop them until they reached the Gulf of Corinth. A band of men was, therefore, sent under the Spartan King, Leonidas, to guard the pass. Only a small army was needed to guard the pass, but unfortunately, a treacherous Greek went to the Persians and showed them another way across the hills so that they could attack the Greeks from behind.

News came to Leonidas that the Persians were advancing and he knew that the Persians were bound to win. So he told his army that those who wished might go away, but that he himself would stay and die fighting the enemy. About 1,000 stayed.

These Spartans seemed unafraid and spent their waiting time by polishing their swords, putting on their best uniforms and combing their hair, so that they might face death with every sign of joy.

As the Persians poured into the plain South of the pass, Leonidas told his men to fight their way out of the Northern end; and there he and his little band died fighting desperately, killing far more Persians than their own numbers. The Persians were astounded at such courage, and angry that so many of their own men had been killed by so few Greeks. Two brothers of Xerxes were killed. Later the Greeks built monuments on the spot where the herioes of Thermopylae had fought, and chief among them was a marble lion to honour the memory of Leonidas.

Question 1: Why did the Greeks oppose (set off to fight) the Persians?

Question 2: Why did the Greeks not run away from the Persian army which was far bigger than their own?

Question 3: The Greeks lost this battle. So why were they ever afterwards proud of their part in it?

The Gunpowder Plot (A Picture History for Boys and Girls, by Richard Wilson).

James I was a King of England and Scotland. He was the son of Mary Queen of Scots. He did not belong to the same Church as his mother; and there some of the people who belonged to his mother's Church were very angry with him because he would not do what they wanted.

A few of these people made a plot. They hired a cellar under the Parliament House where the chief men in the land made the laws.

In this cellar they hid many casks of gunpowder. And it was arranged that a man called Guy Fawkes should fire the powder on a day when the King was to meet the chief men in the Parliament House.

But one of the men in the plot had a friend, Monteagle, who was one of the chief men in the Parliament House. He did not want Monteagle to be blown up so he told him to stay away. Monteagle told the King, who sent soldiers to search the cellars. They found Guy Fawkes and took him away to prison. He would not tell the judge the names of the others in the plot. He was put to death.

Question 1: Do you think it was right or wrong for James not to belong to his mother's Church? Why do you think so?

Question 2 : Was it fair for one plotter to tell his friend Monteagle ? Why do you think so ?

Question 3 : Should Guy Fawkes have told the names of the others to the Judge? Why do you think that ?

THE BATTLE OF CASSINO.

In 1944, when Britain, America and other allies were fighting against the Germans in Italy, the Germans put some of their soldiers in a Monastery which is a place where men live to serve and worship God, on a very steep hill, Mount Cassino, Our soldiers were unable to advance because the enemy could shoot down on them. Several times our men climbed up to drive the enemy out but the hillside was too steep and the monastery walls too high and our men were always driven back.

Then it was said that aeroplanes should be sent to bomb the Monastery. The monks were allowed to get safely out first and then the bombers attacked. After a few bombing attacks the Germans were driven out, but the monastery, which had stood for over 1,000 years, was almost completely destroyed. Our army then pushed on and gradually drove the enemy out of the whole country.

Question 1 : Was it right for the Germans to occupy (capture) the monastery ? Why do you think that ?

Ouestion 2 : Was it right to bomb a house of God. Why do you think that ?

Question 3 : When should the Monks have left their monastery ? Why do you think so ?

IV.—BIBLIOGRAPHY.

[1] CHURCHILL, E. M. (1956). *The Experimental Study of the Growth of Numerical Concepts in a Group of Five-year-olds.* M.Ed. Thesis, Leeds University (unpublished).

[2] CHURCHILL, E. M. (1958). The number concepts of the young child. Leeds University Institute of Education : *Researches and Studies*, Part I, No. 17.

[3] CHURCHILL, E. M. (1958). The number concepts of the young child. Leeds University Institute of Education : *Researches and Studies*, Part II, No. 18.

[4] GUTTMAN, L. (1949). A basis for scaling qualitative data. *American Sociological Review*, Vol. 9, No. 2.

[5] LODWICK, A. R. (1958). *An Investigation of the Question whether the Inferences that Children draw in Learning History correspond to the Stages of Mental Development that Piaoget Postulates.* University of Birmingham Diploma in Education dissertation (unpublished).

[6] PAGE, E. I. (1957). *Haptic Perception.* University of Birmingham Diploma in Education, dissertation (unpubished).

[7] PAGE, E. I. (1959). Haptic perception. *Educational Review*, Vol. 11, No. 2. pp. 115-124.

[8] PIAGET, J. and INHELDER, B. (1956). *The Child's Conception of Space.* Routledge and Kegan Paul.

[9] PIAGET, J., and others (1932). *Moral Judgment of the Child.* Kegan Paul.

A FOLLOW-UP STUDY OF SOME ASPECTS OF THE WORK OF PIAGET AND INHELDER ON THE CHILD'S CONCEPTION OF SPACE

By K. LOVELL

(*University of Leeds Institute of Education.*)

Summary. In their book, *The Child's Conception of Space*, Piaget and Inhelder suggest that a child's first spatial concepts are topological ones and that these later lead to projective and euclidean concepts. A number of experiments from five chapters of their book were undertaken with some 140 children of nursery school age. Part of the evidence presented agrees with that of the authors' and some is at variance with theirs. It is suggested that much more experimentation is needed before the main thesis of Piaget and Inhelder is accepted.

I.—Introduction.

The *Child's Conception of Space*, by J. Piaget and B. Inhelder, is divided into three main sections which deal with Topological* Space, Projective Space, and the transition from Projective to Euclidean Space, respectively. In the opening pages a clear distinction is made between perception on the one hand, and imaginal or representational space on the other. As early as six months of age an ordinary child can distinguish between a circle and a triangle when they are visually presented, but it is much later before he can represent these figures to himself, that is, before he has the concept of the figures. Almost the whole of the book is devoted to the conception of space, not perception as such. The study reported here deals with experiments described in four of the five chapters devoted to Topological Space and the first of the five chapters on Projective Space.

In the opening chapter dealing with Topological Space, the authors deal with 'haptic perception,' or the manner in which the child recognises various objects by touch alone. This is followed by an examination of the drawing of geometrical figures by children, and from the findings, Piaget and Inhelder conclude that in the very early years children construct their space by building up and using certain primitive relationships such as order and enclosure, proximity and separation. These relationships, the authors argue, are topological ones, and they claim that a child's first concepts of space are topological rather than projective or euclidean.

To substantiate their hypothesis Piaget and Inhelder then study linear and circular order through the threading of beads on wires (this involves proximity and separation) in chapter 3 of the first section of their book. This is followed by a number of experiments involving a working knowledge of knots, and in these the child is concerned with surroundings and enclosure. Then, in the last chapter devoted to Topological Space, the authors study the problem of continuity through the gradual subdivision of lines and surfaces into small units and ultimately into points.

Now topological relationships are independent of changes in size of the figure and they do not conserve distance, angles, straight lines, etc., during changes in shape. Thus, topological space allows that kind of analysis which starts from the position of each figure considered in isolation ; it is not a space in which it is possible to co-ordinate all figures within a whole, organised in terms of a common spatial structure throughout. In projective and euclidean space, however, objects are located relative to one another according to perspective or projective systems, or according to co-ordinate axes, respectively. These systems or structures are more complex in organisation and, according to Piaget and Inhelder, are evolved at a later stage in the child's development. Moreover, they argue that projective space begins, psychologically, when the object is no longer considered in isolation but commences to be thought of in relation to a ' point of view.' Thus, in their first chapter devoted to a study of Projective Space, they study projective straight lines and simple perspective.

---

* Topology is the science of that group of spatial relations which remain constant in any spatial transformation ; e.g., open and closed, inside and outside, separation and proximity. Thus, topological relationships are independent of change in size or shape of the figure.

The authors suggest, then, that children progress through a number of rather clear-cut stages at various ages. If this is true, then educationists must take note. For, if they are correct, not only have they shown that the ' space ' of the 3/4-year-old is essentially different from that of the adult, but their work implies that we may have to reconsider the activities involving spatial concepts which our nursery and infant school children engage in, and reconsider our teaching of spatial work in the junior school.

However, when reading Piaget and Inhelder's book, one is struck by two points. First, the authors rarely give exact numbers of children who have been examined or details of the kind of children they are in relation to socio-economic background and intelligence. One suspects there were often only a very few children attempting some of their experiments and these may have come from superior backgrounds. Second, there are often no details of the exact nature of the apparatus used in their experiments, and no indication of the precise phrasing of the questions put to the children—in short no clear standardised test procedure is revealed. Indeed, the present writer had to devise such a procedure as best he could from such indications as are given by the authors.

Another problem relating to Piaget's more recent works must be mentioned as it is possibly of great consequence when considering his work. His early training in zoology, together with his increasing interest in recent years in symbolic logic and the mathematics of groups and sets, may have influenced the model he makes of the genetic structures of the child's mind. Had Piaget had no training in zoology, or had mathematics not reached the stage it has in certain of its branches, one suspects that Piaget might have produced a different ' child.'

Finally, a word must be said about the relationships which are said by Piaget to be topological ones. Are they, in fact, topological relationships at all in the mathematical sense of the word ? Or are young children perceiving and forming concepts involving certain relationships—perhaps somewhat vaguely—which can be put into more precise terms by using topological relationships (which are essentially abstract) in euclidean space ? This is a vital question to which students should give much attention.

Nevertheless, the book under discussion is challenging, stimulating and possibly of great importance, and the writer decided to repeat some of the experiments under more controlled conditions.

II.—PLAN OF THE INVESTIGATION AND STATEMENTS OF RESULTS.

In December, 1957, the Chief Education Officer for Bradford kindly agreed that children in the nursery classes of his Authority could be tested, and the schools were selected by Miss M. L. Beddoes of the Bradford Inspectorate. The head teachers concerned readily co-operated and chose some 150 children between 2 years 11 months and 5 years 8 months of age, varied for age, socio-economic background and general ability as far as could be judged.

It was impossible for the writer himself to examine so many children on account of the time involved. However, the Principal of the Margaret Macmillan Training College at Bradford kindly agreed that twenty-five of her second year students should carry out the experimental work under supervision. Accordingly the writer gave a number of lectures to these students on the views of Piaget and Inhelder. At the same time fourteen complete sets of materials for each of the experiments were made available by the writer and a tentative testing procedure drawn up. The size of the apparatus constructed, and the method of conducting the experiments, were the same as the author's in the main essentials, although, as we have already said, Piaget and Inhelder leave much to the imagination. The experiments were tried out on a group of nursery children at the University of Leeds Institute of Education Children's Centre by Miss E. Steiner, Warden of the Centre. As a result, some changes were made in the wording to be used, and then instructions regarding the testing procedure were duplicated. Finally, Miss Steiner demonstrated the experiments to the twenty-five students with the aid of a few children from the Leeds Centre. The general instructions, a list of materials, and the method used in each experiment are given below. But a description of certain preliminary activities and the instructions given to the children, i.e., a record of the actual words to be used by the students when talking to the children, are not reported in order to save space. These may be obtained from the author. Copies of the proformas on which the children's responses were recorded are not included either. Moreover, it is possible only to give the barest outline of the results.

Each student tested five or six children on seven mornings during the Spring Term of 1958. The students enjoyed this experience and found it another opportunity in which to study children.. Unfortunately, the same children did not complete all the experiments. Owing to illness and bad weather, children were absent from time to time and substitutes replaced the absentees. However, the mean age of the children involved was 4 years 5 months, S.D. 6 months, and these figures varied very little from experiment to experiment, although the numbers of children changed slightly. All the answer sheets were marked by the writer.

GENERAL INSTRUCTIONS

(1) *Surroundings and Attitude.* The best testing situation is that wherein the child feels at ease. Try hard to establish friendly relationships with the child and maintain an encouraging attitude. See the child is comfortably seated and in good light.

(2) *Presence of others.* Try and arrange for no one else to be in the room except another teacher helping you.

(3) *Materials.* Since some of the materials have ' play value ' keep them from view until the actual testing.

(4) *Instructions.* Please keep to the exact instructions laid down. Many people are giving these tests and the manner in which the examining of the children takes place must be standardised.

EXPERIMENT 1—Haptic Perception.

Materials.—A table to hide the objects and shapes when they are being handled. Pencil and paper. Matching charts. Objects and shapes are passed to the child in the *order given below* :

A. 1—a pencil, 2—a key, 3—a comb, 4—a ball, 5—a pair of scissors. A similar set of objects placed within view.

B. The cardboard shapes as follows : 1—Circle, 2—Ellipse, 3—Square, 4—Rectangle, 5—Rhombus, 6—Triangle, 7—Cross, 8—Six-pointed star. 9—Four-pointed star, 10—Cross of Lorraine, 11—Swastika, 12—Semi-circle, 13—Notched semi-circle, 14—Parallelogram, 15—Trapezium, 16—Quadrilateral, 17—Irregular shaped piece with one hole, 18—Irregular shaped piece with two holes, 19—One open ring, 20—One closed ring, 21—Two intertwining rings, 22—Two superimposed rings.

Method.—(i) Child sits at a table, and as he feels each object under the table, he is asked to pick out the corresponding object from among those shown. The objects must be clearly within his view and near enough so that he can point to them.

(ii) Child, while sitting at the table, is handed each of the cardboard shapes in turn. He feels these under the table but does not see them and is asked to pick out the corresponding drawing displayed on the sheet.

Results.

One-hundred-and-forty-five children were tested (seventy-five boys, seventy girls). One mark was given for each correct identification, and the mean score was 8·7, $\sigma=5{\cdot}2$ with no significant sex difference, while the reliability of the test was ·85 as given by the Kuder Richardson full formula. The percentage correct response to the euclidean shapes (Nos. 1–16) was 34·2, to the four euclidean shapes with curved edges (Nos. 1, 2, 12, 13), 50·2, and to the shapes displaying topological relationships (Nos. 17–22) 55·6.

To test for significant differences in response to different kinds of shapes, the ' t ' test was employed. In effect, we tested the hypothesis that, from the point of view of identification, the shapes 1–16 (euclidean) and 17–22 (topological) were random samples drawn from the same population of shapes. The calculated value of ' t ' was 4·2, which for 20 d.f. is significant at the 1 per cent level. Thus, topological shapes were identified more easily than euclidean ones. But, when the identification of topological shapes was compared with that of the four euclidean shapes with curved edges, ' t ' was 1·17. For 8 d.f. this value is not significant at the 5 per cent. level. In other words there is no evidence that the topological shapes are identified more easily than euclidean shapes which have curved edges.

The twenty-two shapes were then ranked for order of difficulty for each of three age groups : up to 4 years 0 months ; 4 years 1 month–4 years 6 months ; 4 years 7 months+, and the rank difference correlation coefficients calculated between age groups. In the youngest group some items were of the same rank, thus distorting the value of ρ and probably reducing its value. Nevertheless, the three values of ρ obtained were ·81, ·73, ·78, suggesting that the relative difficulty of the items, from the point of view of identification, remains roughly the same from age group to age group The five most difficult shapes to identify for the whole group were, square, trapezium, rhombus, parallelogram and quadrilateral. This finding supports the view of the authors that, on the average, straight sides and angles are the most difficult to distinguish by haptic perception.

Piaget and Inhelder (page 21) also state that in their stage 1B (3 years 6 months—4 years 0 months) the square and circle cannot be distinguished because both are closed forms, although they can be distinguished from open forms. Of the forty-three children in our sample within this age range, thirteen, identified both the circle and the ellipse, five identified both the square and circle, three identified both the ellipse and square, and three identified the circle, square and triangle.

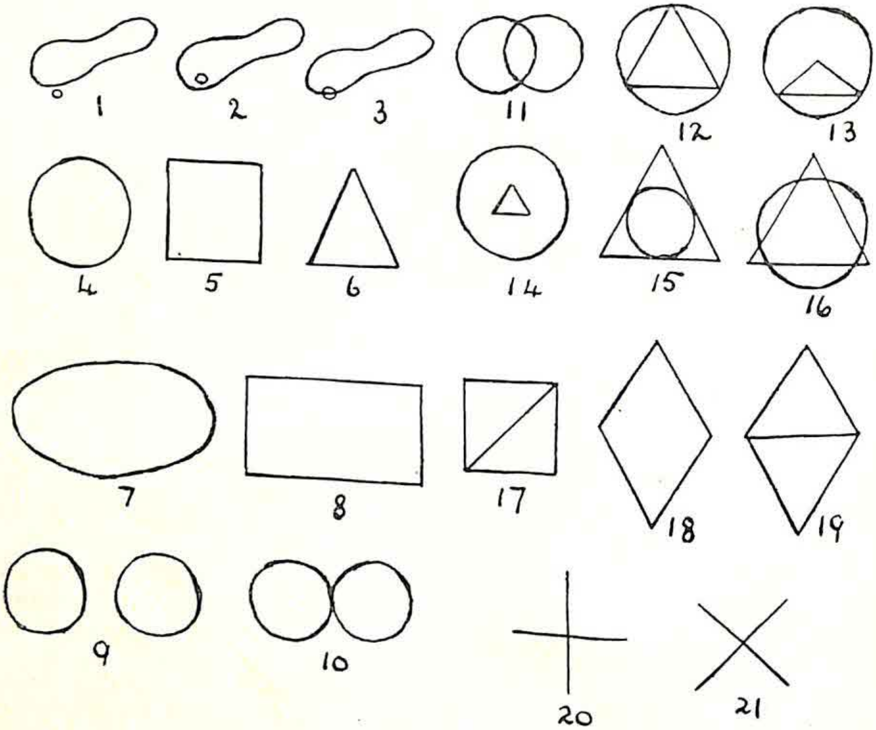
EXPERIMENT 2—The Treatment of Elementary Spatial Relations in Drawing.

Materials.—A set of twenty-one postcards containing the twenty-one shapes given by Piaget. Pieces of paper to match.

Method.—(i) Each child given matches and asked to copy the straight sided figures from the postcards—Nos. 5, 6, 8, 17, 18, 19, 20, 21. Teacher then draws the figure the child makes with sticks on the paper provided, numbering each drawing.

(ii) Child draws a man or woman to get him interested in drawing.

(iii) Each child given a sheet of paper and asked to copy the drawings on each postcard in turn. Have numbers on postcard in top right-hand corner.



Results.

One-hundred-and-forty-five children were involved in this experiment which was divided into two sections, (*a*) the making of straight sided figures with matches, and (*b*) the drawing of twenty-one figures. The marking scheme for both sections is given in the appendix.

(*a*) The mean mark was 8.5, $\sigma=4\cdot2$ and there was no significant sex difference. For the whole group the rank order of difficulty of the figures was No. 5, 20, 21, 6, 8, 17, 18, 19. The rank order of difficulty for the items was then found in each of the three age groups (cf. Experiment 1) and the rank difference correlation coefficients calculated between the age groups. The high values of p obtained, 1·0, ·92, ·92, suggests that the relative difficulty of the item remains almost the same from group to group.

(*b*) The children drew their figures on separate sheets in order to minimise perseveration, which can be common in the age range we tested. All the figures were marked for the presence of euclidean properties in the drawings, while figures 1–3 and 9–16 were also scored for topological properties (e.g., inside, outside, overlap, touching, separation, etc.). The mean score on the euclidean properties was 21·7 (max. possible score 52), $\sigma=12\cdot2$, with no significant sex difference, while the reliability of the test, given by the split half method was ·89. The mean score for topological properties appearing in the drawings was 7·0 (max. possible score 16), $\sigma=3\cdot6$.

Once again the 't' test was employed to determine if the drawing of euclidean properties was significantly more difficult than the drawing of topological properties. That is to say, we tested to see if euclidean and topological properties were random samples drawn from the same population of properties, from the point of view of drawing. The result indicated that, using our marking scale, topological properties were indicated in the drawings more frequently than euclidean properties, ' t ' being significant at the 5 per cent. level. But there was no significant difference between the frequency of appearance of topological properties in figures 1–3 and 9–16, and the appearance of euclidean properties in figures 1, 2 , 3, 4, 7, 9, 10, 11, which involved curves only.

Other points to note are :

(i) The rank difference correlation coefficients calculated from the order of difficulty of the figures in the three age groups were ·75, ·90, ·75, suggesting that the relative difficulty of the items remains fairly constant. The diamond is the most difficult figure to draw with a difficulty value of ·08 for the whole group.

(ii) The rank difference correlation coefficient for the order of difficulty of making figures 5, 6, 8, 17, 18, 19, 20, 21, with matches, and drawing them, was ·93. At the same time, the making of figures with matches was much easier than the drawing of the same figures. This is an important finding and will be referred to again later because it throws some doubt on Piaget and Inhelder's theory of clear stages.

(iii) The product moment correlation cofficient between scores obtained in Experiment 1 and marks obtained in drawing the euclidean properties in the 21 figures in Experiment 2 was ·49. Further, the product moment correlation coefficient between scores obtained in identifying the topological shapes in Experiment 1, and the scores obtained in drawing the topological properties in the figures of Experiment 2 was ·40. There is obviously a good deal of overlap between the abilities involves in these Experiments and it is possible that there is a general ability at work.

(iv) The difficulty values (p) for the drawing of the euclidean properties in the figures for the whole group are :

Figure No.	9	4	20	11	10	5	21	1	2	14	3
p	·73	·63	·62	·60	·58	·48	·47	·46	·46	·45	·43
Figure No.	12	13	8	15	6	17	16	7	19	18	
p	·43	·42	·41	·36	·31	·25	·21	·17	·09	·08	

EXPERIMENT 3—Linear and Circular Order.

Make sure the child understands the idea of colour and that he is not colour-blind. Point to the green bead on the rod and make him pick out one like it. Do similarly with the red and yellow beads. Say in each instance, " Find me one like this."

Section 1.

Materials.—Rod with seven beads on it. Place the rod on the table so that the beads run from left to right, as viewed by the child, in the sequence, Vi, Wh, Gr, Brown, R., Blue, Y.

Method.—Child is asked to place his beads on his rod to make them correspond with the beads on the model.

Section 2.

Materials.—As before plus ' necklace ' of beads.

Method.—Child has to copy the seven beads on his ' necklace ' which is spread on the table before him. Beads are spread fairly evenly around the ' necklace ' and maintained in position with sealing wax. ' Necklace ' spread on table before the child.

Section 3.

Materials.—As in Section 1.

Method.—The child has to put the beads on his rod in the *reverse* order to the beads on the model. The model is placed on the table in front of the child so that the sequence of beads is Vi, Wh, Gr, Brown, R, Blue, Y.

Section 4.

Materials.—' Necklace ' of beads. Rod.

Method.—Teacher makes a figure of eight with ' necklace ' and spreads on the table before the child. Child has to put his beads on his rod in the same order as the beads are on the figure of eight.

Results.

One-hundred and forty-two children were tested. The percentage number of children getting all seven beads correct, also five or six correct, in each exercise, is given in Table 1.

TABLE 1

PERCENTAGE NUMBER OF CHILDREN CORRECTLY PERFORMING VARIOUS TASKS IN EXPERIMENT 3.

	Up to 4 yr. 0 months		4 yr. 1 mo. — 4 yr. 6 mo.		4 yr. 7 mo. — 5 yr. 0 mo.		5 yr. 1 mo. — 5 yr. 8 mo.	
No. of beads in correct sequence....	7	5/6	7	5/6	7	5/6	7	5/6
Copying model direct	15	3	43	20	68	14	78	15
Copying ' necklace ' on to a straight wire	12	9	25	18	46	11	67	7
Making reverse of model	12	0	30	7	51	11	78	4
Copying figure of eight on to a straight wire	6	3	14	2	22	3	56	0
No. of children in each age group....	34		44		37		27	

There was a tendency for the girls to do better than the boys on all the exercises, but the differences were significant only in copying the beads from a ' necklace ' on to a straight wire, and in making a reverse of the model.

EXPERIMENT 4—Study of knots.

Section 1.

Materials.—Overhand knot tightly tied. Ruler or stick.

Method.—Child is shown this knot. He has to make it either entirely on his own, or with the aid of a stick, or after having being shown how to make it.

Section 2.

As for Section 1 plus another overhand knot more loosely formed.

Method.—Child has to follow the slack knot with his fingers, then copy it.

Section 3.

Materials.—Slack overhand knot and a clover knot.

Method.—Child has to follow the clover knot with his fingers.

Section 4.

Materials.—Two pairs of cord circles, one pair intertwined, and the other pair superimposed.

Method.—Child is asked to compare the intertwined and superimposed rings and say what happens to the rings in each pair when they are pulled.

Results.

One-hundred and forty children were examined. Table 2 shows the percentage numbers of children in the age groups who were able to perform the various tasks. More girls than boys were able to make both tight and slack knots, the difference being significant at the 5 per cent. level.

TABLE 2

PERCENTAGE NUMBER OF CHILDREN CORRECTLY PERFORMING VARIOUS TASKS IN EXPERIMENT 4.

	Age			
	Up to 4yr. 0 mo.	4 yr. 1 mo. — 4 yr. 6 mo.	4 yr. 7 mo. — 5 yr. 0 mo.	5 yr. 1 mo. — 5 yr. 8 mo.
SECTION 1.				
Able to make knot	15	48	69	75
Able to make knot with aid of stick	18	0	5	8
Able to make knot after having been shown	39	33	17	13
SECTION 2.				
Able to follow slack knot.	9	36	38	75
Able to copy slack knot	33	52	72	78
SECTION 3.				
Able to follow clover	27	45	38	58
SECTION 4.				
Able to forecast what happens to superimposed rings when pulled	12	29	24	46
Able to forecast what happens to intertwined rings when pulled ..	6	12	19	47
Number of children in each age group	33	42	42	23

EXPERIMENT 5—The Projective Straight Line.

Section 1.

Materials—Ten match sticks each stuck into a small amount of plasticene.

Method.—Two match sticks are set up eighteen inches apart and each 1-inch from the edge of a rectangular table. The child has to place his matchsticks in a straight line between the two end matchsticks.

Section 2.

Materials.—As before.

Method.—The two outside matchsticks are now placed midway along two adjacent sides of the table so that they are about 2 feet or so apart. The child has to place his matchsticks along the straight line forming the two end posts.

Section 3.

Materials.—Round table (hoop) and matchsticks.

Method.—The two outside matchsticks are now placed at the end of a chord of the table—making the distance between them about 2 feet.

Section 4.

Materials.—As in Section 3.

Method.—The sighting posts are arranged as before in Section 3 on the round table. If a round table is not available place the two matchsticks along the adjacent, sides of the rectangular table. In either case spread the remaining ten sticks in zig-zag fashion on the table. The child has to construct a straight line.

In all these experiments the students recorded, by drawing on the specially prepared answer sheets, the positions of the matchsticks.

Results.

Our system of recording the children's responses was, if anything, rather severe, as it is possible that the students made some irregularities when recording the position of the match sticks and these irregularities might have been credited to the children. Table 3 shows the percentage numbers of children who were able to perform various tasks.

TABLE 3

PERCENTAGE NUMBER OF CHILDREN CORRECTLY PERFORMING VARIOUS TASKS IN EXPERIMENT 5.

	Age			
	Up to 4 yr. 0 mo.	4 yr.-1 mo. 4 yr. 6 mo.	4 yr.-7 mo. 5 yr. 0 mo.	5 yr.-1 mo. 5 yr. 8 mo.
SECTION 1.				
Made straight line	53	65	77	80
Slight irregularities in line	40	30	17	20
	93	95	94	100
SECTION 2.				
Made straight line	30	14	32	32
Slight irregularities in line	10	9	20	24
	40	23	52	56
Following edge of table for at least part of the way	40	63	26	36
SECTION 3.				
Made straight line.	20	14	23	36
Slight irregularities in line	17	23	32	28
	37	37	55	64
Following edge of table for at least part of the way	43	47	34	32
SECTION 4.				
Made straight line	10	16	26	16
Slight irregularities in line	20	23	20	44
	30	39	46	60
Number of children in each age group	30	43	35	25

The children's performance over this age range was somewhat irregular. Some children could make a good straight line in Section 2 but not in Sections 3 and 4 and *vice versa*. But there is a consistent improvement in performance with age which is very much in advance of anything Piaget and Inhelder suggest. It is not possible to say why the 4 year 1 month—4 years 6 months age group tended to follow the table edges more. The following points were also noted :

(i) C in the 3 yr. 7mo.—4 yr. 0 mo. age group glances from post to post. So do S and J in the same group although they do not do so on all occasions.

(ii) Some children say that the line is straight even when it clearly follows the bend of the table.

(iii) S in the 4 yr. 1 mo.—4 yr. 6 mo. age group attempts to sight along the line, but not at the level of the table.

(iv) In the 4 yr. 7 mo.—5 yr. 0 mo. age group a few children lay their hands along the table to check whether the line is straight. They also draw imaginary straight lines with their fingers and they are beginning to see the value of standing behind one of the posts.

(v) More use is made of the hands in the oldest age group. There are at least six children in the top age group aiming in one section if not in all.

EXPERIMENT 6.

Section 1.

Materials.—Pencil, circle, rectangle.

Method.—The child has to imagine, and draw, the *apparent* shapes of objects placed in a variety of positions, as ' seen ' by a doll placed at right angles to the child. Then the child has to pick out the correct drawing (from a number of drawings) showing the apparent shapes of the objects.

Section 2.

Method.—Children are asked to draw a picture of straight railway lines disappearing in the distance, and afterwards to add a drawing of the sleepers. Finally, the child has to pick out straight railway lines disappearing into the distance from prepared drawings.

Results.

Piaget and Inhelder (p. 173) state that between 4 and 7 years, on the average, there is a partial or total failure to distinguish between different view points from which the object might be seen. The object is in fact, imagined and shown by and for itself independent of the angle at which the observer actually sees it. In our younger age group we find this to be generally true although in the drawings, position 3, the ' end on ' position, was rather easier than position 2 in which the object was inclined at an angle of 45° to the horizontal. This seems contrary to the authors' findings.

When selecting the drawings our pupils selected correct drawings for positions 2 and 3 more frequently than would happen by chance. Overall, our group is ahead of the author's group in the matter of perspective and it seems as if the child tends to be able to represent to himself the change in the apparent shape of the object before he can draw the change on paper.

IV.—DISCUSSION AND CONCLUSIONS.

Experiment 1.

Children certainly identify, by haptic perception, shapes said by Piaget and Inhelder to display topological relationships more easily than they do euclidean shapes as a whole. However, the authors do not appear to have recognised that euclidean shapes with curved edges are picked out as easily as topological shapes. There is little evidence to suggest that it is topological properties, as such, which enable a child to identify certain shapes more easily than others. It appears that gaps, holes, curves, points, corners, ins and outs, etc., in euclidean space make identification easier. Again, our data do not give much support to the view that up to four years of age children cannot distinguish between a circle, square, ellipse, etc., because these are all closed figures. But our evidence does support very strongly their view that straight sided euclidean shapes with relatively long sides and few corners (square, rhombus, quadrilateral, etc.) are the hardest to identify.

The verbalizations of our subjects were extremely poor compared with those of the author's examinees. When our pupils were asked why they had chosen a certain shape, they invariably replied, " Cos I think so " or a variation of this.

Experiment 2.

Using our marking scheme it is clear that topological properties were better displayed in their drawings than were eucclidean properties taking the figures as a whole. But there was no significant difference between the drawing of topological properties and euclidean properties if we confine our attention to figures which only involve curves. There is clearly a parallel between this finding and that of Experiment 1.

On page 78 the authors say: " there is, nevertheless, a fairly close correspondence between the stages reached in figure construction, irrespective of whether matchsticks or drawings are employed. In only one case, that of a child of 4 who had never previously drawn, were matchstick constructions superior to drawings." Later on the same page they write, " The task of adjusting three matches to form a triangle is just as difficult as the problem of drawing it, and in fact the two problems are solved at exactly the same time." Our evidence does not support this. Table 4 shows that children can construct the figures with matchsticks at least 6 months earlier, on the average, than they can drawn them. While we agree with the authors that the reconstruction of shapes rests upon the actual process of *putting into relation*, it does not seem possible for the child to do this in all ways at the same time. Indeed, our evidence is at variance with the authors' at this point and suggests that more experimental work is required before the latter's theory of clear stages in the development of space concepts from topological to projective ones is accepted. This matter is raised again later in the paper.

TABLE 4

DIFFICULTY VALUES FOR FIGURES MADE WITH MATCHES AND DRAWN IN EXPERIMENT 2

	Square	Triangle	Rectangle	Diamond
Age Up to 4 yr. 0 mo.				
Figure made with matches ……	·67	·35	·45	·13
Figure drawn ………………	·29	·08	·26	·02
4 yr. 1 mo.—4 yr. 6 mo.				
Figure made with matches ……	·68	·66	·57	·23
Figure drawn ………………	·49	·23	·37	·04
4 yr. 7 mo. +				
Figure made with matches ……	·83	·81	·76	·45
Figure drawn ………………	·62	·53	·54	·15

Further, on page 55, Piaget and Inhelder maintain that in sub-stage 1B (3 years 6 months—4 years 0 months) only topological properties are indicated with any degree of accuracy, euclidean properties being completely ignored. Our evidence does not support this. Of the forty-three children aged 4 years 0 months or less, sixteen did draw rather similar shapes for the circle, square and triangle, but a further sixteen produced drawings which clearly distinguished between at least two of these figures. The difficulty values for the figures showed that euclidean properties were by no means ignored in this age group. For example, the square (p= ·29) and the rectangle (p= ·26) are only a little more difficult than figures 1, 2 and 3 which were included because they involved topological relationships, while the circle (p= ·53) is easier than the first three figures. Yet, on page 155, the authors say that the child can only copy squares, rectangles and the like, when he has learnt to copy figures which, while apparently more complex, involve topological relationships only.

Experiment 3.

Reference to Table 1 shows at once that our pupils were more 'advanced' than the sample studied by Piaget and Inhelder. On page 82ff. the authors claim :

(i) Towards the end of Stage 1 (3–4 years) the child can pair off beads on the principle of proximity but he cannot co-ordinate such pairs with one another.

(ii) During Stage 2A (4–5 years as far as can be judged from the book) the child can only make ordered correspondences when the beads are face to face. When there is no visual correspondence the child cannot reproduce order and there is no understanding of 'between.'

(iii) During Stage 2B (5–6 years) a circular order can be transposed to a linear one but reverse order cannot be made.

(iv) Between Stages 2B and 3, at about 6 years, reversed order is gradually obtained by trial and error, but complex pattern-like figures of eight do not give rise to ordered correspondences.

It is not easy to say why our pupils performed so much better than those of the authors. The clear opinion of those who knew the Bradford children we tested was that the average level of their ability was a little below that of a representative sample of British children, and almost certainly below that of the children examined by Piaget and Inhelder. As many as thirty-seven out of the 142 children were able to say that it was the brown bead which remained in position when the model was reversed. Moreover, the authors make no reference to a sex difference in performance in this experiment. This could be a culture pattern effect due to girls playing more with beads and necklaces.

Experiment 4.

On page 107 Piaget and Inhelder say that in the social strata from which their children were drawn, children learn to tie knots around the age of 4, on the average. Our data agrees with this in so far that from 4 years onwards, there is a considerable increase in the number of children who can tie knots without aid. But by 4 years fifteen per cent. could make a knot when asked to do so, almost another one-fifth could make a knot with the aid of a stick, while nearly another two-fifths could make a knot after being shown.

A comparison of Table 2 with the findings of the authors shows that correspondence appears to a child when the perceptual pattern varies earlier than they would have us believe. In Experiment 3 it is a question of discontinuous correspondence, i.e., the correspondence of individual beads. In this experiment we have correspondence which is continuous on two counts. There is correspondence affecting both order and surrounding of sections of cord which are continuously linked together.

Experiment 5.

This is a very important experiment. If children can make a straight line using matchsticks (placed up-right) before they can draw a line when attempting the side of a square or triangle, then the theory that the child's first concepts of space are topological ones leading to concepts of projective space cannot be accepted on the evidence available at the moment.

For Piaget and Inhelder the construction of a straight line may be defined as the linking of distant points by the interpolation of a series of points along a direct path. They claim that the following stages can be identified :

(i) Stage 1—up to 4 years. " The child is equally incapable of forming a straight line—even when it runs parallel to the edge of the table—as of drawing the straight sides of a square or triangle given a model to copy " (p. 158).

(ii) Stage 2 (4–7 years). The child can form a straight line more or less correctly when it runs parallel to the edge of the table, but is unable to do so when it lies at an angle to the sides of the table except at the end of the period. Up to 6 years of age the child cannot break away from the perceptual influence of the table ; between 6 and 7 he slowly overcomes this difficulty by trial and error (p. 158).

(iii) On page 160 they repeat that up to 6 years of age the child has no clear idea of what constitutes a straight line in the sense of having a symbolic image going outside the field of perception, or as regards organising a new construction within the framework provided by perception. When it is a matter of forming straight lines with matches, even parallel to the edge of the table and only 2/3 cm. from the edge, they fail utterly and completely. We quote, " Their performance is confined simply to the formation of a topological line with successive elements very close together and curved in various ways."

Our evidence indicates that in our sample some 53 per cent. in the youngest group made a good straight line and a further 40 per cent. made only slight irregularities in it, when asked to make a line parallel to the edge of the table. With the oldest age group, the corresponding percentages were 80 and 20 respectively. Again in Sections 2, 3 and 4 the performance was far better than the authors would lead us to expect, although perceptual influences were clearly at work in some children, as can be seen by examining Table 3.

Indeed, we do not agree with Piaget and Inhelder that the child up to 4 years 0 months is as incapable of making a straight line parallel to the edge of the table—or at an angle to the table edge or as a chord of a circle—as of drawing the straight sides of a square or triangle given the model to copy. If we include lines with slight irregularities with the good straight lines—this is entirely fair as the child was not penalised for slight irregularities when drawing a straight line—then the difficulty values for the various tasks are, for the youngest age group :

Making a straight line parallel to the edge of the table	·93
Making a straight line at an angle to the edge of the table ..	·40
Making a straight line as a chord of a circle	·37
Drawing a square	·29
Drawing a triangle	·08

The most frequently used principles, as far as could be ascertained, in lining up the matchsticks were those of working from post to post, of working from the middle outwards, or from outside posts inwards. This finding was also in evidence in an experiment carried out by one of the writer's students, Mr. D. F. Clarke, A.T.D., using a small group of very retarded children aged 5–11 years. Some children look from post to post but do not aim as such, although a few children do take aim—more in the older age groups. The authors are generally right in stating that aiming does not take place on a widespread scale before 6, but they are wrong in suggesting that straight lines are not made.

In conclusion, we can say that our evidence sometimes agrees with that of Piaget and Inhelder and sometimes it is at variance. The variability in performance within an age group is far greater than they allow. In the writer's view, the crucial question is one bearing on the author's main theory. Does the child's conception of space begin with topological concepts which are transformed concurrently into concepts of projective and euclidean space ? The data which we obtained, expecially in Experiments 2 and 5 suggests that this should remain an open question until further experiments have been carried out. Perhaps the child's world is not so different from that of the adult as Professor Piaget would have us believe. The present writer is not the first person to suggest this.

ACKNOWLEDGEMENTS.—The writer takes this opportunity of thanking Miss M. F. Brown and Miss H. M. Edwards, Vice-Principal and Senior Lecturer in Education respectively, at the Margaret Macmillan College, for the help they gave in this experiment. He is also greatly indebted to the students, to Mr. Spalding, the Chief Education Officer for Bradford, to Miss M. L. Beddoes, and to the head teachers of the schools concerned. A research grant from the University of Leeds Institute of Education covered the expenses which were involved.

V.—APPENDIX.

SCORING OF TEST 2.

Arrangements of Matches.—Two marks for each correct arrangement. One mark for not more than 1 error in orientation *or* arrangement.

Drawings.—(*a*) Euclidean concepts. Two marks for a correct drawing. One mark if not more than one error, or two errors symmetrically executed; e.g., rectangle drawn as a square.

In the case of shapes involving curves, two marks if fairly true to shape, one mark if it can be clearly recognised as say a circle or ellipse.

But, in the case of Figures 12, 13, 14, 15, 16, four marks for euclidean properties i.e., two for each figure. In Figure 16 in particular, one mark each for the Triangle and Circle of both approximately the same size; but one mark only if triangle well within circle or *vice versa*.

(*b*) Topological concepts. In Nos. 1-3, 9-16, one mark for topological property shown (i.e., separate from marks for euclidean concepts); e.g., one figure inside another.

Thus, in No. 10, two marks if circles drawn under E. concepts, one further mark separately scored if circles touching. If one shape inside another but shapes quite incorrect then no marks for E. but 1 for T. concepts.

In Figures 12, 13, 15, two marks for T. properties, i.e., one for enclosure, one for touching. In 16 three marks for T. properties.

(*Manuscript received, June 27th*, 1958.)

SYMPOSIUM: THE DEVELOPMENT OF MORAL VALUES IN CHILDREN

PUNISHMENT, DISCIPLINE AND EDUCATIONAL OBJECTIVES

BY D. K. WHEELER

(*Faculty of Education, University of Western Australia.*)

SUMMARY.—The nature and extent of punishments administered in a representative sample of all Western Australian secondary schools are reported and supplemented by information about the punishment of these same adolescents in their own homes. The reactions to punishment of 765 students are analysed and discussed. Most students conceive of punishment as retaliatory and/or retributive. This is taken to be a reflection of the community attitude towards punishment, which is contrasted with modern preventive and remedial theories of discipline. The effects of sanctions commonly used in schools are considered with respect to educational aims in a democratic society, and the conclusion drawn that such negative sanctions do not effectively promote, and may even retard, the educational process.

I.—INTRODUCTION.

No arguments can possibly arise about the need for order and discipline in social situations, particularly in the school classroom, where they are vitally necessary to the learning process. It would seem, however, that concepts of discipline, methods of organization conducive to order as a necessary condition of learning, and types of sanctions common in some schools, exhibit marked disparity from those which might be deemed most effective by the social scientist.

This article reports the nature and extent of, and reactions to, punishment administered to adolescents in Western Australian secondary schools. After consideration of the experimental material, and comparison with the nature of and attitudes to punishment administered in the home, it discusses the lack of change in school sanctions over a long period of time and the effectiveness of discipline based upon negative sanctions in preparing adolescents for life in modern democratic communities.

II.—THE EXPERIMENTAL INVESTIGATION.

In 1951 the Faculty of Education of the University of Western Australia embarked upon an extensive investigation of adolescents in secondary schools.* A sixteen-page questionnaire was administered to a sample stratified on the basis of:

(*a*) sex composition of the population at school;
(*b*) proportion of rural to urban schools;
(*c*) school type—(i) Government, (ii) Roman Catholic, (iii) Private or Independent (non-Government, non-Roman Catholic).

The sample consisted of 290 first-year secondary school students (148 boys, 142 girls), 289 third-year (136 boys, 153 girls) and 186 fifth-year (113 boys, 73 girls) at school in 1951 (comprising 4 per cent. 10 per cent. and 24 per cent. of the school population for each year). Mean ages for these three levels were

---

* This was organized and carried out by Julie M. Wheeler, who worked in a part-time capacity as Senior Research Fellow from 1951-53, under a grant from the Research Grants Committee.

13 years 5 months, 15 years 4 months, and 17 years 2 months for boys, with girls a month younger in years 1 and 3.

Amongst the many questions asked were the following :

124. Are puuishments given in your school ? 125. What sort ?
126. Do you think it is right for this sort of punishment to be given in school ?

TABLE 1

PERCENTAGES OF RESPONDENTS MENTIONING VARIOUS TYPES OF PUNISHMENT IN SCHOOL.

Type of Punishment	Boys						Girls						
	Non-Government		Government				Non-Government			Government			
	Roman Cath.	Inde-pend-ent	J.H.S. * and R.T.	3 yr. H.S.	5 yr. H.S.	Tech.	Roman Cath.	Inde-pend-ent	Bus. Coll.	J.H.S. * and R.T.	3 yr. H.S.	5 yr. H.S.	Tech.
Corporal	90	74	79	97	—	—	16†	—	—	40†	3†	—	—
Impositions	34	36	24	12	25	—	44	37	17	25	62	26	12
Detention	1	47	3	2	—	—	24	77	22	8	18	3	25
Deprivation	3	1	29	5	5	—	4	12	—	36	6	14	6
Demerit marks	—	—	—	—	—	—	—	—	—	—	—	—	—
Silence	—	—	—	—	—	—	23	3	—	—	—	—	—
Sent out	—	—	6	—	8	—	—	—	—	3	3	6	—
Sent to H.M.	—	—	—	—	4	—	—	2	5	—	14	9	—
" Lectured "	—	—	3	2	12	10	4	—	—	—	1	9	12
Fatigues	—	19	—	3	—	—	—	33	—	—	—	—	—
None	18‡	5‡	—	1	57	85	6	—	55	3	10	56	56
No. of Respondents	79	77	34	111	76	20	70	57	18	36	95	76	16
No. of Mentions	100	138	49	133	41	19	70	94	8	26	99	52	18

* Junior high schools and rural tops (i.e., primary schools with some secondary classes).
† Refers either to primary school or to boys.
‡ All these were fifth year boys who said that corporal punishment is used only in lower forms.

III.—ANALYSIS OF RESULTS.

The detailed analysis of the responses is shown in Table 1. The data were analysed by individual schools, sex and years, but for convenience have been reported by types of schools. As some respondents mentioned more than one type of punishment, percentages in the columns sometimes add up to more than 100. The nature of the categories is self-evident, though Deprivation is rather a blanket term covering loss of privileges or denial of participation in sporting or other school activities.

(*a*) *Nature and Incidence of Punishment.*

From the responses, Western Australian schools fall into two classes, those in which the official policy is one of no punishment, and those in which formal punishments occur. In the first group fall the Government five-year high schools and the technical schools, and the business colleges. Formal punishments occur in all other Government schools, in Roman Catholic schools and in independent schools. Over half the students in five-year high schools and business colleges say there is no punishment and 85 per cent. of boys and 56 per cent. of girls in technical schools say the same thing. Apparently the customary policy is not to have formal punishment (except, as some students point out, expulsion) though individual teachers occasionally depart from this.

Punishments common to almost all schools are impositions, detention and deprivation of privileges. Another commonly imposed is caning or strapping of boys. Demerit marks and fatigues are confined to certain types of schools, while being sent out of the classroom, or to the principal, and verbal castigation are also used. Some of the answers would also indicate that verbal castigation may accompany other types of punishment.

Certain types of punishment appear to be used in certain schools, but not to any extent in others. For boys, caning or strapping is common in all schools, except Government five-year high schools and technical schools.* Impositions are used in all schools, but detention and fatigues only in independent schools in a systematic way. Deprivation of privileges is most common in rural tops and junior high schools.

For girls, impositions are again common, though one suspects that in business colleges (where 55 per cent. of respondents say there is no punishment) ' extra work ' means doing the original work correctly. Detention is also common though most used in independent schools. Deprivation of privileges (usually sport) is most common in junior high schools and rural tops, though it occurs less frequently in other schools. Demerit or order marks appear in Roman Catholic schools and fatigues in independent schools. Being sent out of the classroom is reported only from Government schools, though one respondent from an independent boys' school said, ' The master walks out ' (but this is apparently not a common punishment). Verbal reproof is more often reported from Government schools.

Certain entries in Table 1 appear to offer contradictions, in that over half the respondents from certain types of schools say there is no punishment, while the remainder detail various kinds of punishment they undergo. Experience in some of these schools indicates, as was suggested earlier, that school policy is as indicated by the majority of the students, but that individual teachers do not always conform with this policy. The 18 per cent. of boys from Roman Catholic schools and the 5 per cent. from independent schools who say there is no punishment are specifically referring to themselves as seniors who are apparently rarely punished. In schools where corporal punishment is used, over 30 per cent. of the fifth-year boys specifically state that such castigation is always or mostly reserved for lower forms and does not apply to the last two years.

---

* Not all Independent Boys' Schools use corporal punishment.

The 16 per cent. of girls in Roman Catholic schools and the 40 per cent. or so girls in Government schools who report corporal punishment in their schools refer to younger children in primary departments or to boys in co-educational schools, though there are one or two secondary school girls who speak feelingly of the use of rulers and of slaps.

There are no significant differences between the sexes or between the years in frequency of mention of punishment. Although the two younger groups of boys tend to make mention of more punishments than the girls, it would be unwise to equate frequency of mention with frequency of punishment or to draw any conclusions about whether one sex is punished more than another.

(*b*) *Attitudes to Punishment.*

Boys: The boys' attitude to punishment administered in schools seems a reflection of the attitudes in the community that make such punishment possible. These 397 boys refer to various forms of punishment, the total number of mentions being 465. Of these, 314 say " Yes, it is right for this sort of punishment to be given in school," 106 say " No," and 45 give a qualified answer, such as, " It's right if you deserve it, but the right person doesn't always get it." In general, the attitude of those who approve of the status quo is, " How else could we be punished for wrong doing ? " or " It teaches us not to do wrong," or, " How else could be children kept in order ? " There is some difference in the attitudes of the first and later years. (first and third years $\chi^2 = 8{\cdot}55$, 2 d.f.; $p = {\cdot}02$; first and fifth years $\chi^2 = 12{\cdot}05$; $p = {\cdot}01$). As these boys grow older, more of them tend to approve of the punishment given, fewer disapprove and rather more tend to give qualified answers. It may be that the older boys receive less punishment and hence approve of it for others, or it may be that they hold even more firmly to the community values.

This change does not appear with regard to corporal punishment, which receives 262 mentions. Of these 172 express approval, 68 disapproval and 22 give a qualified answer. There is no significant diffierence in attitude between any two years. If anything, the trend is for more of the older boys to express approval. Certainly, a number of the fifth year boys expressly stated that corporal punishment was not administered to them, now that they were seniors, but still said that they thought it right that it should be used in school. Typical of these who said ' Yes ' were the following remarks : " Teaches you to be better," " They (the teachers ?) have to have law and order somehow,' " It's the only effective sort." Those who did not think it should be used said " Mediaeval corporal punishment," " The cane does not have much effect on high school students," " It's completely illogical." Qualified answers were of the type " What else could you use ?," " If you deserve it," or " If it's not too severe."

With respect to other forms of punishment, the first-year boys did seem to differ from the older ones. (Years 1 and 3, $\chi^2 = 6{\cdot}16$, $p = {\cdot}05$; Years 1 and 5, $\chi^2 = 8{\cdot}33$, $p = {\cdot}02$). Once again the older ones give a straight-out ' Yes ' more often and ' No ' less frequently, though qualified answers appear more often in the fifth year. Suggestions made above may well apply here.

Girls.: The older group are significantly different from the two younger groups in their attitude towards punishment. (Years 1 and 5, $\chi^2 = 12{\cdot}33$, $p = {\cdot}01$; Years 3 and 5, $\chi^2 = 11{\cdot}27$, $p = {\cdot}01$). The only suggestions that can be made here are those already offered, that the older girls approve of punishment more because they are less subject to it (that they approve of others being punished) or that they have taken over the values of their culture and reflect them in their responses. Certainly, these older girls do not specifically state that they are exempted from punishment as do the senior boys with regard to corporal punishment, but experience in schools with both boys and girls leads one to believe that the incidence of punishment is considerably less in the senior year. In general, as one student put it in connection with a school without formal punishment, " Your are expected to act as adults."

(c) *Punishment in the Home.*

It has been suggested above that the attitude of these boys and girls towards punishment is a reflection of the community's attitude, and further light is thrown upon this by adolescent answers to two other questions which were asked earler in the questionnaire. " What sort of punishment do you usually receive at home? " and " What do you think about it? "

The great majority of adolescents received one of the following forms: (i) *Corporal* (beating with strap or cane, hitting, smacking, ' belting '); (ii) *Deprivation* of some pleasure, recreation, outing or privilege; (iii) *Verbal castigation* (' growling at,' ' ear-bashing,' ' telling-off '). Among 13-year-olds, significantly more boys receive corporal punishment, while significantly more girls are deprived of pleasures or lectured. One half of the boys at this age receive corporal punishment, but it is surprising to learn that one-fifth of the girls are disciplined entirely or in part in this manner.

By the age of 15, reports of corporal punishment diminish considerably, and deprivation and lecturing become the most common forms, the earlier sex differences disappearing. This decline in corporal punishment in the home, especially of boys, is marked and significant, and it seems as though bodily chastisement is replaced by verbal reproof. Is it that boys become more amenable to discipline, or do they transgress less; or is it simply that their size and physical development make corporal punishment less possible? It may even be possible that the decline is due to the composition of the sample and that parents who allow or insist on schooling beyond the minimum age of fourteen are those who tend most to eschew corporal punishment at all ages.

By the age of 17, corporal punishment is practically non-existent in the home, or is confined to a slap or box on the ears. Lecturing is the most favoured sanction, then deprivation, or else there is no punishment. In this oldest group, parents apparently resort to a process of reasoning or lecturing or verbal attack (' ticking-off '). Only one-fifth of parents of this age group rely on the adolescents' own reactions to the incident, i.e., employ no sanctions.

When they were asked their opinion about the punishment administered by their parents, these adolescents were free to express themselves as they felt inclined, and most of them did so. Their replies fall into the following categories:

(i) *Punishment is an effective deterrent, good character training.*—Typical replies are: " good for me," " stops you from playing up," " I have to be taught right and wrong," " Does me good, if I didn't get the strap I would grow up a ruffian."

(ii) *Punishment is justified retaliation for wrong doing.*—Respondents in this class seemed to believe that punishment for wrong doing is deserved, that parents have the right to punish erring children, that it is necessary and has to be done.

(iii) *Punishment is sometimes justified.*—These adolescents apparently feel the same about punishment as those in category 2, but believe that their parents are not always fair, i.e., that sometimes the children do not 'deserve' the punishment.

(iv) *Punishment as an unpleasant process.*—The replies in this group stress strongly the 'awfulness' or painfulness of punishment, mostly with respect to corporal punishment. Typical answers were "has me in tears," "terrible," "horrible," "cruel."

(v) *Indifferent to Punishment.*—These respondents claimed that they did not mind their punishment, and that it did not make any difference to their behaviour. In endeavouring to give the impression that punishment had no effect on their behaviour, they were possibly expressing their resentment and defiance.

(vi) *Punishment is wrong or unfair.*—A few expressed their belief that punishment was 'not right.'

(vii) *Punishment not severe enough.*—A few cases felt that they were getting off too lightly for their own good and expressed a wish for more severe punishment.

On the whole, the 13-year-olds approve of all types of punishment. Boys regard corporal punishment as retaliation for wrong doing (and as unpleasant in itself), for it was clear from the protocols that parents administered sound beatings, not token slaps. The girls were more divided in opinion about physical castigation, fewer regarding it as either right or justifiable. Punishment was regarded as good character training or as an effective spur to future good conduct by about 20 per cent. In general, these 13-year-olds tend to view punishment as a form of vengeance, which they accept when it is felt to be closely related to their naughtiness.

Surprisingly enough, the incidence of punishment for 15-year-olds is not much less than for the younger group and again the majority approve or accept what they receive. Corporal punishment is accepted as a necessary evil, while deprivation and verbal reproof are now more likely to be considered as right and effective rather than mere retaliation for wrong doing. It is clear that those who are not punished achieve this happy state because of strict obedience and exemplary behaviour, not because their parents disapprove of punishment or because the respondents are thought mature enough to do without it.

Even at the 17-year-old level, half the boys and 70 per cent. of the girls consider that punishment is good for them, or is good character training, or that it is an active deterrent from wrong behaviour.

IV.—DISCUSSION.

It is clear from this material that adolescents' attitudes towards punishment in school are reflections of their attitudes towards punishment in the family, and one may suppose that family practices tend to reflect the climate of opinion of Western Australian society. What stand out most clearly are the following ideas:

(*a*) The 'right' of adults to punish children and young people.
(*b*) Punishment is retaliation for wrong doing.
(*c*) Punishment is a deterrent from wrong-doing.

Modern educators, like modern penologists, no longer subscribe to vindictive, retributive or deterrent theories of punishment, but tend towards remedial and preventive theories. This, of course, does not mean that individuals are excused from responsibility for their actions. Diagnosis and cure are the desired processes in *remedial* discipline: in *preventive* discipline, the central idea is that, if right conditions exist, misbehaviour is at a minimum.

There is, then, apparently considerable disparity between the views of the modern educator and those of members of this community, including the teachers in it. While individual teachers may subscribe to remedial or preventive theories, many would seem to hold with retributive or deterrent ones. Further evidence that can be produced in support of this view of the community attitude is found in the Education Regulations, 1949. As Government schools cater for three-quarters of the school population (70 per cent. at the secondary level), and as other schools show little or no difference in the actual teaching situation, we shall not go far astray if we consider only this particular authority for the sanctions employed.

Only two punishments are mentioned in the Education Regulations, 1949 (Regs. 80-82 and Reg. 180), namely corporal punishment and detention. Corporal punishment can be administered (under certain conditions) for offences against morality, gross impertinence, and wilful and persistent disobedience, but not for failure or inability to learn, for not preparing home lessons, or for trivial breaches of discipline. Detention (of individuals) may be used for idleness, unpunctuality, disobedience, or similar faults,* but not for inability to learn.

Cetainly, over the past fifty years there has been a change in educational concepts of discipline and order, but this has not touched every school in all respects. Lancaster (1805) lists as punishable offences: Idleness, talking (repeated or frequent offences after admonition), disobedience to teachers or monitors, profinity, slovenliness and offences against morality. With the exception of talking, those are almost precisely the same as the offences for which punsihment is laid down in the Western Australian Education Department Regulations, 1949. Moreover, the relevant Regulations are almost literal transcripts of regulations issued in 1907. It looks as though the school norms (the formalized procedures and modes of intereaction) have not changed greatly in fifty or one-hundred-and-fifty years. This appears scarcely consistent with current advances in knowledge about motivation, learning, child development or the dynamics of groups.†

It may be asked what is the effect of these various punishments upon the individual in the learning situation. In Government schools the severest punishment (corporal) is reserved for offences against morality and the teacher's authority, unless consistently poor performance in school work can be interpreted as persistence and wilful disobedience. (Detention is not extensively used in these schools, perhaps because " it is often needful the master should confine himself in school."‡) Much has been written on corporal punishment, so that all that can be offered here is the statement that there is little to recommend it. Schools have shown that it is possible to educate without corporal punishment.

---

* Whatever this may mean. What faults are similar to idleness, unpunctuality and disobedience? Though this terminology offers difficulties in classification, presumably the teacher who detains children finds it a useful ' catch-all ' provision.

† Yet even Lancaster is tolerably ' modern ' in some of his statements. Cf. " If any misconduct should be punished by severity, vice and immorality are the chief subjects; and, I am convinced that it *is not always indispensable in those cases*, having known many a sensible boy reformed without, and that, from practices as bad as almost any that usually occur in schools." (p. 100.)

‡ Lancaster, *ibid.*, pp. 103-104. " Few Punishments are so effectual as confinement after school hours. It is, however, attended with one unpleasant circumstance. In order to confine the bad boys in the schoolroom after school hours, it is often needful the master, or some proper substitute for him, should confine himself in school to keep them in order. This inconvenience may be avoided by tying them to the desks, in such a manner that they cannot untie themselves."

The onus of proof that such schools are in any way inferior must fall upon those who believe in the cane or the strap.

All schools use impositions extensively and all subjects pay their dues in this field. Doing impositions in a subject is recognized as punishment by the children : how far does doing similar work in the ordinary course of the school day come to be regarded in the same manner. After all, we know that attitudes *are* engendered in behavioural situations. Replies from these students suggest that impositions are probably the most common punishment for infractions of school discipline and for inadequate behaviour in the learning situation. Logically and psychologically, it might be reasonable to suppose that their effectiveness is inherent less as punishment than as deterrents to interest in and liking for the particular subject chosen as an imposition.

On the face of it there does not seem any great harm in detention, unless the detention period is filled in with impositions. In such cases, the previous remarks apply. Nor does deprivation seem so educationally undesirable except that often the student is debarred from participation in some vital part of the curriculum, usually sporting activities.

Both observation of the community and the general results of this investigation would indicate that in the socialization and education of children, negative sanctions are used more than positive ones. Yet Hurlock (1924) and others have shown that motivation, success and reward are more effective in learning situations than punishment. Adults in learning situations expect reward rather than punishment, and, if they do not meet with success, tend to " leave the field." It might be well for schoolmasters to consider the proportion of rewarding remarks and behaviour which occur in classrooms, compared with reproof and punishment.

Too many parents and teachers still place great faith in rules and regulations, in demanding unquestioning respect for authority, obedience without question, freedom from annoyance for themselves, and still cling to retaliatory or deterrent theories. Modern educators have abandoned these ideas and, like Burton (1944), believe that " Pupils do not learn responsibility for conduct through undergoing bullying, through enduring sarcasm, through suffering harsh punishment, through the inflexible enforcement of rules and regulations. They learn responsibility for conduct through being shown the reasons for good conduct, through being treated as honest, sincere individuals who will co-operate if given a chance and through being given increasing opportunity for exercising decision and selt- control."

We believe with Burton that the teacher's general aim in the maintenance of order should be the development within the group of a social conscience and group co-operation, and the achievement by the individual of social ideals, attitudes and habits which make possible self-control and responsibility. We also believe that many teachers have insufficient knowledge of principles and techniques which enable them to work towards these ends. So that, in desperation, they fall back on punishment as a deterrent or as retribution. Keeping in mind the fallacy of confusing number of mentions with incidence of punishment, we still believe from the answers given and our own educational experience, that far too many teachers use negative formal sanctions (punishment) where diagnosis of the situation and proper application of knowledge of child development and social psychology would better serve educational ends. This is essentially a matter of competence. Many teachers are competent in their subject matter but incompetent in their knowledge and practice of individual and group psychology, and, what is as much to be deplored, in their knowledge of the findings of experimental education.

In an efficient learning situation the social processes of indoctrination, habituation, utility and group identification work towards the teacher's ends and are aided by positive sanctions, so that punishment is unnecessary. It is only when curricular offerings and teaching methods are unsuitable, or where the norms of the classroom group differ from those of the school (as in classes from sub-cultures) that there is the possibility of conflict between school norms and those of student groups. Results from other parts of the adolescent survey show that most of the adolescents accept the cultural norms with respect to conduct and that nowhere were there indications that even boys could achieve status by rebelling against adult authority as represented by school or parents. Group standards are so strong in adolescence that it seems a pity that the strength of "what the group thinks" is not utilized more by the teacher in the attainment of adequate educational ends. There are many areas in the school where students could profit from the opportunity to learn to make decisions, rather than having to comply with teacher decisions under penalty of negative sanctions.

If the final outcome of school discipline should be understanding and responsibility manifested in changed behaviour, that is, the internalization of social and moral codes rather than outward conformity with and control by rules, then we must seriously consider whether in the school we set up a morality of constraint or a morality of co-operation, to use Piaget's (1932) terms. If we wish, in our classrooms, to foster co-operation and the development of true morality as distinct from the morality of constraint, we shall try to modify the authoritarian, teacher-dominated competitive situation (Cf. Wheeler, 1957) where the student's sole responsiblity is to see that he does what he is told and that he does not transgress the rules, lest he be punished. One of the developmental tasks (Havighurst, 1953) that all children must face, and one that is particularly important in the adolescent period is the achievement of a mature set of values and a set of ethical controls characteristic of a good man and good citizen. As teachers, we must consider this particular developmental task and the experiences offered secondary school students to help them tackle it, as well as the proper place of external controls (including punishment) in learning situations.

Hemming (1957) has suggested that ways in which the school can affect the process of acquiring values are by reducing the influence of values in the community which are felt to be undesirable; by promoting the child's self respect; by educating his moral insight. That many members of the community hold retributive and deterrent theories of punishment inconsistent with democratic ideals does not necessarily mean that professional educators should also subscribe to them: and what better way of promoting the child's self respect than by slowly replacing a morality of constraint with one of co-operation? Revision of some of our ideas about punishment may lead to the education of moral insight. The problem, as Piaget puts it, is to know what will best prepare the child for his future task of citizenship. "Is it the habit of external discipline gained under the influence of unilateral respect and of adult restraint or is it the habit of internal discipline of mutal respect and of 'self government'?" We should think carefully about the part students may play in school government.

Both the evidence from this questionnaire and experience in schools would indicate that Western Australian secondary schools are still largely authoritarian in nature. Without elaboration, we suggest that this condition arises from the following factors:

(*a*) Uncertainty about the aims and objectives of the school in a period of rapid cultural change.

(*b*) A curriculum once perhaps valid for the secondary school population, but now unsuited to a majority of students. In this connection, it should be noted that a committee of educationists and laymen has been working on this problem for some time.

(*c*) Insufficient application to the educational sphere of existing knowledge of child development, social psychology, experimental education and the dynamics of group behaviour.

(*d*) Continued use of a structural and almost invariably authoritarian organization perhaps useful in earlier times but certainly more relevant to earlier political theories than to the working out of a successful democracy.

(*e*) Continued reliance on deterrent and retaliatory theories of punishment instead of acceptance of remedial and preventive ones.

(*f*) Lack of a diagnostic approach to questions of order and discipline.

It seems clear that if we want a working democracy we must train for it in our schools. Example is considerably more potent than exhortation and the acquisition of mere verbal knowledge. We must give children opportunities to learn to govern their affairs (including their learning) and themselves. This is no plea for soft pedagogy, for essentially the acceptance of responsibility for one's own actions is a more difficult task than the acceptance of external authority. Students can learn to be responsible only by being given training and experience in taking responsibilities within their capacity. Paternalism in schools, however, beneveloent, must give way to distributive leadership (cf. Gibb, *et al.*, 1951) and external authority, at least to some extent, be replaced by the authority of mutuality. School administrators and teachers must carefully scrutinize the basis of authority and examine the kindred concepts of leadership and competence to see how all three can best be used to achieve the ends of the school. Above all the diagnostic approach to the learning situation (of which order and discipline are concomitants) must be the teacher's professional attitude. Diagnosis is a fundamental professional technique of the doctor ; so it should be of the teacher.

V.—REFERENCES.

[1] BURTON, W. H. (1944). *The Guidance of Learning Activities.* New York : Appleton-Century-Crofts.

[2] GIBB, J. R., PLATTS, N., and MILLER, F. (1951). *Dynamics of Participative Groups.* Chicago : John S. Swift Co.

[3] HAVIGHURST, R. J. (1953). *Human Development and Education.* New York : Longmans.

[4] HEMMING, J. (1957). Some aspects of moral development in a changing society. *Brit. J. Educ. Psychol.*, 27, 77-88.

[5] HURLOCK, E. B. (1924). The value of praise and reproof as incentives for children. *Arch. Psychol.*, 11, No. 71.

[6] LANCASTER, J. (1805). *Improvements in Education as it Respects the Industrious Classes of the Community.* London : Darton and Harvey.

[7] PIAGET, J. (1932). *The Moral Judgment of the Child.* London : Kegan Paul, Trench, Trubner and Co.

[8] WHEELER, D. K. (1957). Education and group dynamics. *Educand*, 3, 69-80.

[9] EDUCATION DEPARTMENT OF WESTERN AUSTRALIA. *The Education Regulations*, 1949. Perth : Govt. Printer.

(*Manuscript received July 21st*, 1958.)

TRENDS IN EDUCATIONAL PSYCHOLOGY *

BY STEPHEN WISEMAN
(School of Education, University of Manchester.)

I.—INTRODUCTION.

MY purpose in this paper is to try to survey the trends in educational psychology over the last few decades. In any field of research one is likely to find a certain degree of imbalance ; an over-emphasis on problems which, at the time, appear acute, and yet which, in the end, may prove little more than blind alleys ; and a corresponding neglect of other aspects of the subject which may need active investigation before progress in other sections of the field become possible. It is also possible that outside pressures may force us to concentrate on applied research—the attempted solution of acute contemporary problems—when a more sober and far-sighted judgment might advocate a sustained attack on the more fundamental matters subsumed under the category of pure research.

II.—SCOPE OF RESEARCH.

The first fact to come to light in a review of this kind is the large increase in the amount of research devoted to educational psychology over the last forty years. If we consider the number of theses presented for higher degrees in education in the United Kingdom (listed in Mrs. Blackwell's excellent series of books published from 1950 onwards under the auspices of the National Foundation for Educational Research) we find the average number per year rising from 23·5 in the decade 1918-1927, to 137·5 in the period 1948–1955, as will be seen from Table I :

TABLE 1

THESES PRESENTED FOR HIGHER DEGREES IN EDUCATION IN THE UNITED KINGDOM (FROM BLACKWELL'S LISTS).

Period	Average No. of Theses Per Year	Percentage dealing with Psychology and Method
1918–1927	23·5	49·0
1928–1937	67·0	59·1
1938–1947	64·0	63·6
1948–1955	137·5	76·2

The percentages of these concerned with psychology and method (as against philosophy and history) are seen to rise from 49 per cent. to 76 per cent in the same period. The total volume of research in any particular field is, by itself, not necessarily highly significant. Quality is perhaps more important than quantity, particularly in research. And we all know how possible it is for research—and particularly, perhaps, research in education—to be badly designed, indifferently carried out, or focussed on unimportant problems. There is, too, the contrast between what I have called ' pure ' research and the more practical ' applied ' research. The teacher in the classroom may often sigh for immediate answers to day-to-day problems, and reasonably so ; yet the teacher of teachers perhaps hopes more for an extension of theory as a result of research, a comprehensive theory which will form the basis for generalized advice and teaching. The attack on single problems without concern for basic theory is a wasteful process, and one which can advance our knowledge only slowly and almost accidentally. In the first stages of a 'new' subject such research is inevitable; only when usable theories become to be framed, and are then used to guide and inform further research, can we feel that the subject is growing-up.

---

* Paper given at Section J, British Association, Glasgow Meeting, August 29th, 1958.

Here are two quotations which we may contrast. "Reasons for many of the deficiencies of educational research are found in the immediate needs of application and the inability to defer them while developing a proper theoretical framework." Nevertheless, another writer says: "But in child psychology we are still faced with the danger of theory being exalted above fact." Curiously enough, the first quotation is contemporary (1957), and the second quotation was written in 1931.* Perhaps at the earlier stage the too-facile development of theory on an inadequate basis of research might well have merited the stricture.

There seems little doubt that, at the present time, too much research is being done of a fragmentary nature, and too little is designed for the isolation of crucial problems and the development of comprehensive theory. There are many reasons for this; some of them are considered later in this paper. Too often it seems that educational psychologists are prepared to wait for other workers to produce their theoretical frameworks for them. Let us take, for example, the psychology of learning. Textbooks on learning theory have been produced in which children are never mentioned. The pages are full of rats—with, perhaps an occasional pigeon or octopus. The animal psychologists are largely responsible for the research on which the learning theories of Hull and Tolman are based. I am not, of course, complaining because animal psychologists are 'muscling in' on our own territory: we must, indeed, be profoundly grateful for the massive attack which has been mounted in this way. What I do deplore is what seems to be a relative lack of interest and of participation of educational psychologists in this work. As the biologists, ornithologists and animal psychologists pursue their inquiries into the problem-solving of the lower forms of life, cannot the educational psychologists open a second front—perhaps through the study of children's abilities? We are all familiar with the considerable progress made in this field under the stimulus of the practical problems of personnel selection and of secondary school selection, and we recognize the fundamental achievements of Burt and Vernon in this country in mapping out for us the now familiar hierarchical structure of abilities. Is it not possible, however, that, by orientating our efforts towards learning theory rather than the solution of practical problems, by studying the child's capacity to learn through his emergent abilities, then we may develop this second prong of the attack? Another trend in this field is in the psychology of personality, in which many 'pure' psychologists are considering personality as the product of the pressures and tensions of the environment. How the individual reacts to these, how he responds to different kinds of stimulus, how he 'learns' to behave in particular situations, may be explicable in terms of learning theory. So while the educationist is trying to produce learning in the classroom, and is occasionally investigating isolated learning situations, the animal psychologists and the clinical psychologists are developing a theory of learning which should be his primary concern. This is a bad thing from two points of view: first because it is a serious criticism of the professional competence of the educational psychologist; second because the emergent learning theory might be less productive, less comprehensive and slower in growing, without the help of critical studies in the field of child development carried out by experts.

---

* KELLER and CORCORAN (1951). *Ann. Rev. of Psychol.*, p. 177; HAZLITT (1931). *Brit. J. Educ. Psychol.*, 1, p. 127.

That this is felt also in the United States is clear from Stroud's complaint in the 1951 *Annual Review*: "... one of the pressing needs in our field is a comprehensive learning theory (or behaviour theory) that will encompass teaching and learning in school and embrace the broad phenomena of personality and adjustment." In the 1952 volume, Elmgren re-iterated: "The need for a meaningful theory of learning that really influences the teacher . . . still persists . . . Experimental learning psychologists in the U.S. as elsewhere have almost completely neglected the study of children's learning in the school situation."

III.—THE LAST TWENTY-FIVE YEARS.

For the purposes of this paper, I have endeavoured to survey in some detail the last twenty-five years. In making this survey I have relied upon a number of sources. First is the coverage of papers in the *British Journal of Educational Psychology*, the first number of which appeared in 1931. The American opposite number—the *Journal of Educational Psychology*—has also been used. Such journals as these are extremely useful for survey purposes, but they have disadvantages. The main difficulty is the fact that what appears in a journal may be the result of editorial policy to a very large extent. The fact that, for example about 30 per cent. of the articles in the *B.J.E.P.* in its first ten years of life, were concerned with the psychology of learning and teaching, cannot be taken to show that 30 per cent. of the educational research in Britain was being done in this field. Many papers (particularly in the *B.J.E.P.*) are solicited by the editor. Professor Valentine, who edited the *British Journal of Educational Psychology* so brilliantly until 1956, made a point of publishing symposia on subjects which he felt were important but neglected. In this way trends may be obscured—or, indeed, new ones stimulated.

For the more recent past, the *Annual Review of Psychology*, and particularly the chapters on child psychology and educational psychology, have been invaluable. This excellent publication did not appear until 1950, unfortunately, but since then it has provided an extraordinarily good coverage.

These sources have been used, of course, to provide a check and a supplement to the writer's own impressions of developments in the subject. Such checks are very necessary, since one's own specific interests and research activities inevitably lead to a lop-sided view of the field.

Let us look at some of the data showing where the emphasis lies in research. The analysis of the coverage of the *British Journal of Educational Psychology* and of the (American) *Journal of Educational Psychology* has been supplemented from two other sources. In the 1955 volume of the *Annual Review* the authors of the chapter on Child Psychology produced an analysis of the 185 research references they cited. In 1953 the U.S. Department of Health, Education and Welfare produced a Bulletin, *Research Relating to Children*, which classified research in progress. These two latter sources are not immediately comparable with the journal analyses, since they deal with 'child psychology' and not 'educational psychology'; they do not, therefore, include research on teachers and their problems, nor on specific problems of teaching method. Nevertheless, the common ground is considerable.

Four categories have been chosen under which to classify the topics reviewed:

(1) *Learning*: Including the teaching and learning of particular school subjects; the psychology of learning and teaching; growth, development and school progress.

(2) *Abilities*: Including the theory of abilities; measurement techniques; individual differences; selection for secondary education.

(3) *Emotional Development*: Including adjustment; motivation; interests and attitudes; social behaviour; personality dynamics.

(4) *Teachers*: Professional problems; personal adjustment; selection and training.

Such a classification is, of course, quite arbitrary, and the allocation of a particular article or research topic is often a highly subjective judgment. Nevertheless, it is a rough and ready guide with which to explore the field.

Table 2 shows the analysis of the two journals of educational psychology.

TABLE 2

CLASSIFICATION OF JOURNAL ARTICLES (PERCENTAGES).

	Brit. J. Educ. Psychol.			J. Educ. Psychol.*	
	1931-40	1941-50	1951-57	1930	1953-54
1. Learning, etc.	31	24	37	22	31
2. Abilities, etc.	43	40	44	62	40
3. Emot. Dev.	16	27	12	16	27
4. Teachers	10	10	8	0	2

* Based on D. G. Ryans, 1955, *Ann. Rev. Psychol.*, 6, p. 432.

Notice how the first category—learning—although showing a slight upward trend, still only represents one-third of the articles published. And this is true on both sides of the Atlantic. Category 2 (Abilities) is perhaps a little misleading for the British journal, since it includes selection for secondary education, a topic which has been of prime importance in England since 1944. The periods 1941-50 and 1951-57 saw the publication of a considerable number of papers in this field, including a long symposium. If these papers are taken out, the percentages for 'Abilities' then show no rising trend, as will be seen from Table 3.

TABLE 3

CLASSIFICATION OF BRITISH JOURNAL ARTICLES AND DEGREE THESES (PERCENTAGES).

	Brit. J. Educ. Psychol.*			Blackwell's Lists
	1931–40	1941–50	1951–57	1952–55
1. Learning, etc.	33	27	44	35
2. Abilities, etc.	39	32	32	27
3. Emot. Dev.	17	30	14	33
4. Teachers	11	11	10	5

* Omitting articles on 11-plus selection.

The field of personality, adjustment and emotional development shows a rise from 16 per cent. to 27 per cent. in America over a twenty-five year period. The British rise over the first and second decades is not continued in 1951-57, but in this perhaps the journal does not truly reflect the overall trend in research. The classification of the theses in Blackwell's most recent lists is given in Table 3, and from this it seems highly probable that the relatively low percentage of journal articles in this category shows the effect of the editor and his policy. In America the stress on personality and emotional development is even stronger—as might be expected, since it often seems that 'fashions' in psychology tend to transfer from the U.S. to Britain with an appropriate time lag. Table 4 reviews research in progress in America, and this reveals a strong emphasis on emotional development, adjustment, social behaviour, etc. Very much more than half the total research falls within this category.

TABLE 4

REVIEW OF U.S. RESEARCH IN PROGRESS (PERCENTAGES).

	Ann. Rev. of Psychol., 1955*	U.S. Dept of Health., 1953†
1. Learning, etc.	29	23
2. Abilities, etc.	12	13
3. Emot. Dev. etc.	60	65

* Adapted from Radke-Yarrow and Yarrow, 1955, *Ann. Rev. Psychol.*, p. 3.
† Bulletin " Research Relating to Children."

It is significant, I think, that a figure of 60–65 per cent. from these two sources is reduced to 27 per cent. for the American *J.E.P.*, and that 33 per cent. from Blackwell's lists is reduced to 14 per cent. for the *British J.E.P.* Can it be that over half the papers submitted to the editors *do* fall in this category and that the radical reduction in the proportion accepted for publication is caused, not only by editorial policy, but by the poor technical level of many of the research papers? From my own experience it would seem true that there are many contributions of this kind which are fragmentary and anecdotal, lacking in experimental rigour and unsupported by any coherent basic theory.*

IV.—A PRESENT TREND.

The complaint that the lag of theory is acting as a brake on progress is even more true in the general field of child development than in the field of learning.

---

* Some editorial comment seems to be called for, though I cannot, of course, speak for Professor Valentine who was Editor up till the end of 1955. I would very much doubt whether the editorial policy of this *Journal* in the 1950's has been unfavourable to articles in the field of Emotional Development, and would suggest that the variations to which Dr. Wiseman calls attention may well arise: (*a*) from mere chance fluctuations over shortish periods, (*b*) from difficulties of categorisation. Thus, when I attempted to classify articles from 1951-7 under Dr. Wiseman's headings, I arrived at 30 per cent. on Emotional Development, instead of his 14 per cent.! Admittedly, however, a colleague who undertook the same task independently agreed more closely with him, reaching a figure of 17 per cent.

I have also compared articles accepted with those rejected from 1956 to 1959. The percentages concerned with Emotional Development (according to my scoring) are 32 per cent. of acceptances and 39 per cent. of rejections. This difference of 7 per cent. is not statistically significant and, therefore, does not support the hypothesis that the Editor is rejecting articles in this field appreciably more often than in other fields. P.E.V.

We have studies of ' behaviour ' rather than studies of children. ' Development ' has become separated from ' learning.' The emphasis is very strongly on social interaction and ' problems ' of behaviour. In the United States and in the United Kingdom it has become unfashionable for child psychologists to consider the intellectual growth of children. Only with animals does it seem respectable to analyse problem-solving. R. G. Barker, in 1951, commented " A great superstructure of applied clinical psychology appears to be in the process of erection with little actual concern for the foundations that are said to be laid in childhood."* It is probable that no adequate theory of development is possible without a preceding theory of behaviour, which itself must be integrated with, and consist largely of, a comprehensive learning theory. The theories of development which seem to underlie the work of some child psychologists are little better than those which were current in training colleges one or two generations ago. Many of us are familiar with the ' race-recapitulation ' hypothesis, which compared the development of the child with the development of the race. This was an interesting speculation, and one which was admirably adapted as the basis of a lively student essay, but in its usual form it was of no value as a *theory*. It did nothing to explain *why* children did certain things at certain times, but merely gave a spurious respectability to the assertion ' a child does so-and-so because he is a child ! ' Other theories of the ' stage ' and ' phase ' type, emphasising the ' years of dependence,' or the ' gang age ' similarly do little more than describe observations on some children at some times. Whether the samples were adequate ; whether all children go through such stages, or only some ; whether the limits of the ' stage ' are well-defined and in what terms ; such things are rarely, if ever, considered. But what is mainly lacking is the attempt to develop a theory to explain those observations which are themselves masquerading as a ' theory.'

There is clearly a trend at the moment in this particular section of educational psychology, and the figures given above support this view. The existence of the trend is understandable. In the '30's we grossly neglected the social aspect of learning and development. We felt that child growth and education was mainly a matter of cognition ; general intelligence and special aptitudes were the crucial factors. The pendulum has swung pretty far since then. Now the emphasis is on social inter-action, on sociometry and sociology, and by the time the teacher has coped with ' pressures and tensions and needs in the highly-structured classroom field ' she has precious little time to spend on teaching the academic subjects. Let me again give two quotations from educational psychologists, one from America, one from Scotland : " One cannot escape the impression that the accent to-day is on child growth and development. Academic subjects and skills are not at the focus of educational discussion." And again : " Though the chalk legends on school walls that AB loves CD have not yet attained the status of educational research, some of the contributions . . . especially the British, are only about one stage beyond this."† My concern here is not to denigrate educational sociology or sociometry ; I am not suggesting that ' social interaction ' is unimportant ; and I am well aware that some research in these fields is competent, rigorous and highly significant. What I am deploring is the over-emphasis on this side, and the corresponding neglect of the intellectual aspects of learning and teaching. In the field of child development we must study the interactions of many factors—in particular intellectual endowment, maturation, intellectual exercise and social exercise. It is impossible to study interactions by isolating any one of these factors.

---

* *Ann. Rev. of Psychol.* 2, p 20.

† *Ann. Rev. Psychol.*, 2, 283 (Stroud) and 5, p. 365 (Maxwell).

It is true that these trends in Britain and America are not universal. In Geneva, for example, we have Piaget working away on child development with methods and theories strikingly different from the current research I have quoted. His publications, however, have caused more stir among the 'general' and the 'pure' psychologists than among the educational psychologists, who seem more prone to follow fashion and less receptive to new methods and ideas than are their colleagues. Piaget's success in developing a creative school in child psychology is due to his brilliance as a clinical psychologist and to his insistence on developing theory and hypothesis, based on past results, to guide further experimentation. The clinical approach, however, brings its own dangers. Too much of Piaget's theory is based on results of a handful of children. What is now needed is the re-iteration, in Britain and America and elsewhere, of his crucial experiments, using careful sampling and well-designed experiments.

V.—THE FUNCTIONS OF RESEARCH.

Up to this point research has been considered only in its major aspect of providing answers to problems and of furthering our knowledge of children and of teaching. This is not its only function, however. Almost as important—in the long run—is its educative influence on the researcher. A good deal of research (perhaps the greater part of it) is done by teachers, and the value of this to the teacher is very considerable. Planning, executing and reporting an investigation, under skilled guidance, is an educational process of a unique kind, leading to new insights, changed attitudes and greater professional competence. The usual way in which a teacher comes to research is by way of a master's degree in education. The research topics which are possible for a master's thesis tend to be limited in scope and restricted in range—naturally enough—and so the amount of growth in knowledge from such research findings is relatively small. The chief function of the master's degree is the training of the students in research methods and techniques. Those teachers who so show themselves to be competent research workers must then be encouraged to do further—and more productive—research. This would normally be done through a Ph.D. degree. But, when we analyse the situation in Britain, we find a deplorable state of affairs. From 1918 to 1951 there were 1,663 master's degrees awarded in education, and 374 Ph.Ds. Of these doctorates, 223 were conferred by London, and a further 81 by the Scottish Universities.* During this period 560 master's degrees were awarded at the Universities of Birmingham, Bristol, Durham, Leeds and Manchester; in the same period only eleven Ph.Ds. were given.† In other words, the provincial universities (excluding London) produced one in three masters' degrees, and only one in thirty-four Ph.Ds. Thus, our provincial universities train teachers to do research, and then, having trained them, prevent them from going further. This 'prevention' occurs because the Ph.D. regulations are framed primarily with science students in mind, and demand one or two years' of full-time research. Now the teacher's laboratory is his classroom, 'part-time' Ph.Ds. in education are long over-due.

One of the striking facts about educational research is the small number of training college lecturers who are involved in it. And yet, there are many reasons why such lecturers ought to do research. They are able to work in many schools within an area, and can, therefore, tackle much needed research on teaching method. Few teachers can do this because of the problems of administration and co-operation involved. They can work in nursery, infant and junior schools; most university researchers naturally choose the secondary school. Most lecturers are themselves highly qualified academically and have proved themselves very good teachers; a high proportion ought to be capable of doing good research. They themselves need the education of research training in order to be better trainers of teachers. And, most of all, the colleges where our teachers are being trained ought to be enlivened and stimulated by fostering research programmes within their walls. If our student-teachers do not come into contact with research during their training, most of them never will. Can we wonder that it takes so long for research findings to be translated into classroom practice when we neglect to educate our teachers in their attitude to research? Is it surprising that teachers' organisations—many of them extremely wealthy—rarely consider spending money on research? Education must be the only existing field where research is hardly ever financed by the *users* of research.*

---

* While it is true that London is more accommodating in admitting External and part-time Ph.D. students, it should be pointed out that three-quarters of all *internal* Ph.Ds. in educational psychology at London are obtained by students from overseas. Most of the remaining quarter of candidates are on the staffs of training colleges. Thus, it is almost as rare in London as in the rest of the country for a practising school teacher to obtain a Ph.D. P.E.V.

† WISEMAN, S. (1953). Higher degrees in education in British universities. *Brit. J. Educ. Studies*, 2, 54-66.

The deficiencies of educational research are not caused by 'the immediate needs of application and the inability to defer them while developing a proper theoretical framework.' This may be so in the United States: in the United Kingdom it is probably due to the lack of interest —or awareness—of local authorities and teachers' organisations; the slowness of the universities in adapting their regulations to encourage trained research workers; the over-full programmes of training college lecturers and the lack of appreciation that teachers ought to be trained in institutions which have active research programmes going on. It is a sad fact that the teaching of educational psychology in training colleges too often falls into one of two categories. There is first the 'traditionalist' method: students are taught what the lecturer was taught twenty years (or more) previously. McDougall's instincts and emotions are still the last word in some of our colleges. Then we have those on the 'lunatic fringe' who eagerly embrace the latest fashions, and attempt to teach to students ill-digested 'facts' and 'theories.' The *Journal of the Association of Colleges and Departments of Education* recently produced a number specially devoted to the teaching of educational psychology which contains some examples of the second category.

VI.—CONCLUSION.

It will now be obvious that what I hope for above all is the development of a really comprehensive theory of children's learning; and that, based on this, there will arise a more adequate theory of child development. This can only happen, I believe, if the present trend towards more and more studies in the social psychology of the classroom is slowed down, and the emphasis shifted towards cognition and 'academic' learning. This is more likely to be achieved within the next two or three decades if the financing and direction of research is widened in scope; if our universities make it more possible for teachers to do research at the doctorate level; and if our training colleges insist on research as an indispensable element in the activities of their staff so that teachers in training are brought into contact with the growing edge of education.

---

* It should be recognised, of course, that most L.E.As. and most national associations of teachers give financial support to the National Foundation for Educational Research: subscriptions from these sources in 1957-58 were nearly £14,000 from the former, and £2,300 from the latter. In addition to this, many L.E.As. pay the tuition fees of teachers doing research for the M.A.(Ed.) and M.Ed. degrees: it is doubtful, however, whether this would amount in total to more than a few hundred pounds per year. Not only are these sums inadequate, but there is also the danger that authorities might feel the payment of a *per capita* subscription to the N.F.E.R. effectively discharges all their research reponsibilities.

THE STANDARDISATION AND PREDICTIVE VALUE OF A BATTERY OF ACADEMIC ACHIEVEMENT TESTS FOR THE PURPOSE OF ACADEMIC COUNSELLING AT SOUTH AFRICAN UNIVERSITIES*

BY M. C. H. SONNEKUS

(Faculty of Education, University of Pretoria, South Africa.)

SUMMARY. In this investigation an attempt was made to construct and standardise academic achievement tests for the purpose of counselling at South African universities.

Tests in Afrikaans, English, mathematics, physics and chemistry for first year university students were combined in two batteries for the prediction of academic success in B.A. and B.Sc. courses respectively.

The results proved that the tests standardised in this research are better predictors of academic success than intelligence tests, ability (aptitude), interest, and personality tests.

A comparison of these test results with those in overseas researches shows that they compare favourably.

I.—INTRODUCTION.

A.—*The Problem Stated.*

IT is well known that numerous problems present themselves to the first-year university student who finds himself or herself confronted with the problem of becoming a member of a new community. Apart from the more general problems of self-adjustment, social adjustment, accommodation, finances, study habits and study methods, he is concerned about his academic progress. Further difficulties often occur later on account of an unmotivated or wrong choice of subjects.

Important indications of academic maladjustments are, for example, the number of failures at university and the number of students who abandon their studies before completion. Nel and Gouws[9] showed that in South Africa the percentage of first-year students failing at South African universities varies between 20 per cent. and 40 per cent. In a preliminary report dealing with the study results of students at the Rijks University Utrecht, during the period of 1946-1954, Rümke[11] shows that about 30 per cent. of the students at this university abandon their studies before completion. The investigations of Wenberg[14] at the University of Minnesota revealed that in 1946 and 1947 20 per cent. and 32 per cent. of the students respectively, withdrew from the university and gave as first and principal reason academic problems, i.e., difficulty with choice of subjects.

Considering the number of directions of study and the diversity of subjects within these different directions, from which the inexperienced first-year university students must choose, the problem of the correct choice becomes more pronounced. Such students also experience a certain amount of uncertainty in their choice of a future vocation and reveal a certain degree of ignorance of vocational conditions. In this respect Webb,[13] of the Department of Psychology at the University of Washington, found that 52·6 per cent. of the men students and 61·1 per cent. of the women students had not yet decided on their vocation or were still unsure about their choice. This uncertainty results in a large number of changes in the choice of subjects during the first two months of the first year courses.

---

* This survey is an abbreviation of a dissertation submitted by the author for the degree of doctor in the Faculty of Education, University of Pretoria, South Africa. The author wishes to express his gratitude to Professor B. F. Nel, Dean of the Faculty, under whose guidance the investigation was carried out. A word of thanks is also due to the Research and Publications Committee, University of Pretoria, for a grant which made this research possible.

It is patently clear from these and numerous other investigations that first year students need academic counselling in the choice of subjects. The nature and extent of this academic counselling fall outside the scope of this study. The following are a few aids which may be used to facilitate the choice of subjects: Intelligence tests, achievements in the matriculation examination, the cumulative school record card, personality tests, interest tests and student questionnaires.

An aid which is not yet used at South African universities is academic achievement tests, i.e., standardised achievement tests in specified academic subjects. These tests yield information about the achievement level of the first-year university student in specific academic subjects and are already largely used in some overseas countries, especially the U.S.A., to provide academic counselling for first-year university students in the choice of subjects.

B.—The Purpose of the Investigation.

The purpose of this investigation is then to standardise academic achievement tests which may be used to aid the first-year student in choosing those subjects that are adapted to his ability and interest.

The battery of tests covers the following subjects: Afrikaans; English; mathematics; physics; and chemistry. The intention is to combine the five achievement tests into two batteries relevant to the choice of a B.A. or B.Sc course respectively. The first-year university student should be examined at the university by means of these tests before selecting his subjects, thus obtaining information concerning himself, which will facilitate his correct choice of subjects.

C.—Method of Investigation.

(1) In each of the five subjects mentioned, a preliminary academic achievement test was drawn up in accordance with specific academic requirements for matriculation standard.

(2) These preliminary academic tests were applied to first-year university students from a number of Afrikaans-speaking universities. The investigation was not extended to English-speaking candidates. All students had studied these subjects to the end of the matriculation course at the secondary school.

(3) Following an item analysis of the preliminary tests, the final tests were drafted.

(4) The final tests were applied to a second group of first-year university students.

(5) The validity and reliability of each of these tests were calculated by utilising the final results obtained at the end of the first year.

(6) The tests were combined into two batteries with the aid of multiple correlation and the calculation of Beta weights, thus enabling predictions to be made of academic success in a B.A. or B.Sc. course of study.

II.—THE EXPERIMENTAL INVESTIGATION.

A.—The Preliminary Standardisation.

After a preliminary test was drafted for each of the above-mentioned subjects the tests were applied for the first time at the beginning of 1956 to first-year university students at Pretoria, Potchefstroom and Bloemfontein (all three Afrikaans-speaking universities). The numbers ranged from 500 taking Physics to 811 taking English. Preliminary time limits were set for each test.

An item analysis was then carried out, first to show the degree of difficulty of each of the test items.[1] Subsequently the discrimination values were calculated for all items by utilising an internal criterion. For this purpose tetrachoric correlation coefficients were used. The test distribution was partitioned at the median, and from the percentages of testees in the upper half and in the lower half who answered an item correctly, the correlation was read off from an abac graph.[8]

Thereafter, the degree of difficulty as well as discrimination values were used to select the test items and thus to draft the final test. All test items correctly answered by from 40 per cent. to 60 per cent. of the test persons were retained, provided they also showed good discrimination values. As far as possible only test items with discrimination values above ·20 or ·30 were included. Items were re-arranged in order of difficulty.

Separate answer sheets were drafted and the preliminary time limits were adjusted to appropriate lengths for the final tests. In scoring, the usual guessing correction for wrong answers was applied.

B.—Final Standardisation.

At the beginning of 1957 the final tests were applied to similar numbers of first-year university students at the afore-mentioned South African universities, as in the previous trial.

Some modifications were made in the time limits to ensure that at least 80 per cent. completed each test. These times are shown in Table 1. From the test distributions, percentile norms were calculated.

C.—The Determination of the Reliability of the Tests.

For this purpose only the results of the first-year university students of the University of Pretoria were used. Split-half coefficients corrected by means of the Spearman-Brown formula are shown in Table 1.

TABLE 1

RELIABILITY COEFFICIENTS OF ACADEMIC ACHIEVEMENT TESTS.

Academic Achievement Test	Reliability coefficient	Number of persons tested (N)	Number of test items in the test	Time Limits in minutes
Physics	·907	474	120	45
Chemistry	·827	464	139	45
Mathematics	·811	513	58	58
Afrikaans	·659	624	65	20
English	·832	614	60	20

A comparison with the reliability coefficients of overseas academic achievement tests indicates that, with the exception of the Afrikaans test, the reliabilities are generally satisfactory. Possibly the Afrikaans test is unduly short, and it is in any case the first test of its kind. The reliability of the Co-operative School and College Ability tests of the Co-operative Test Division, Educational

Testing Service[4] varies from ·82 to ·95, and of the Iowa Placement Examinations[3] from ·87 to ·98.

D.—The Validity of the Academic Achievement tests and the Combination into Two Batteries for the Prediction of Academic Success.

(1) *The question of a criterion for the determining of academic success.* Seeing that academic achievement tests have as their aim the prediction of academic success, the criteria used throughout this investigation were the final results of the first-year students as obtained in examinations at the end of the first year.

It was necessary to find a technique of converting marks of different examinations to a common scale which does not necessarily require that the marks in each subject display a normal distribution ; for in some subjects there were fewer than thirty candidates. A method used by Gouws[5] and the Student Guidance Committee of the University of Pretoria was applied.

The distribution for any subject was divided into ten (not necessarily equal) class intervals, so that each class interval included as nearly as possible, one-tenth of the total number of first-year university students who had taken it. These decile groups, from 1 to 10, were then used as the candidates' scores.

Note that the procedure assumes that the differences in the distributions of marks between different courses is a function of examination techniques and standards, and does not represent a difference in basic potentiality. In other words the groups of students who took the different subjects were assumed to have been drawn from the same universe with regard to basic academic potentiality. Although this assumption is not necessarily correct, it is more acceptable than the hypothesis that the original distributions of marks in the different subjects were drawn from the same universe.

(2) *Correlations between academic achievement tests and total final achievements (decile scores) of first-year university students (compound criterion).* Table 2 shows the correlations between each test and the total deciles for 5 B.A. (occasionally 6) and 4 B.Sc. subjects.

TABLE 2

CORRELATIONS BETWEEN ACADEMIC ACHIEVEMENT TESTS AND TOTAL FINAL ACHIEVEMENTS (COMPOUND CRITERION).

Academic Achievement Tests	B.A.—Total	N	B.Sc.—Total	N
Afrikaans	0·245	217	0·303	236
English	0·399	218	0·278	246
Mathematics	0·237	141	0·485	235
Physics	0·311	118	0·448	212
Chemistry	0·333	122	0·418	215

All correlations significant at the 1 per cent. level.

Conclusions.—From Table 2 it can be concluded that the tests are valid predictors of academic success in the B.A., as well as the B.Sc. direction of study, the validity coefficients for B.Sc., being the highest in general.

In comparison with overseas tests, the Scholastic Aptitude Test[10] of the College Entrance Examination Board displays validity coefficients which vary from ·21 to ·52, and the Co-operative School and College Ability Tests[4] of the Educational Testing Service, give coefficients of ·32 to ·60. Generally, the validity coefficients in this investigation, compare well with overseas tests.

Table 3 shows correlations obtained for intelligence tests.

TABLE 3

Correlations Obtained for Intelligence Tests* (Compound Criterion).

Field of Study	Intelligence Tests		
	U.S. Group Intelligence Test	South African Group Intelligence Test	New S.A. Group Intelligence Test
B.A.	·321* ·303*	·393*	·295*
B.Sc.	·322* ·101	·164	·261†

* Significant at the 1 per cent. level.
† Significant at the 5 per cent. level.

Conclusions.—(i) For B.A. courses the correlations obtained from the tests in this investigation (see Table 2) compare well with those obtained with intelligence tests, with a slight difference in favour of the academic achievement tests. (ii) For B.Sc. courses the academic achievement tests are better predictors of academic success than intelligence tests.

Table 4 shows the correlations obtained for the Differential Ability Test Battery of the National Bureau of Educational and Social Research (South Africa).

Conclusion.—It appears that the five academic achievement tests with a few exceptions (compare Table 2) are better predictors of academic success for B.A. than the Bureau tests, while for B.Sc. they are by far better predictors than the Bureau tests.

In further research by the Student Guidance Committee, low insignificant and in some cases negative correlations were found between the Kuder Interest Questionnaire and marks in B.A. courses (B.Sc. courses could not be included due to a lack of sufficient data); and for the Bernreuter Personality Questionnaire, low insignificant correlations of ·024 and ·071 for B.A. and B.Sc. respectively, were obtained.

In comparison with these results, it is clear that the five academic achievement tests in this research, with significant validity coefficients for B.A. of ·237 to ·399, and for B.Sc. of ·278 to ·485, are far better predictors of academic success.

---

* The University of Pretoria appointed a Student Guidance Committee in October, 1952, to investigate the nature, the causes and extent of maladjustment of first-year university students. Findings in Tables 3, 4 and 8 concerning intelligence and ability (aptitude) tests, were obtained from an unpublished report by this committee.[12]

TABLE 4

CORRELATIONS OBTAINED FOR THE DIFFERENTIAL ABILITY TEST BATTERY OF THE BUREAU OF EDUCATIONAL AND SOCIAL RESEARCH (COMPOUND CRITERION).

Bureau Tests				
B.A.		B.Sc.		Tests
1953	1954	1953	1954	
·332*	·297*	·179	·033	Afrikaans comprehension.
·291*	·116	·238†	·155	Afrikaans vocabulary.
·272*	·422*	·050	—·236†	English comprehension.
·272*	·320*	—·110	—·052	English vocabulary.
·163	·251*	·166	·056	Memory.
·380*	·227*	·378*	·313*	Verbal reasoning.
·147	·260*	·261†	·156	Non-verbal reasoning.
·034	·096	·119	·020	Space perception.
·087	·273*	·218	—·030	Arithmetic (mechanical).
·093	·214*	·271†	·138	Arithmetic (problem).
·010	·045	—·180	·232	Mechanical comprehension.

* Significant at 1 per cent. level.
† Significant at 5 per cent. level.

From the comparisons so far, it seems that of *all standardised tests academic achievement tests are the best predictors of academic success.*

(3) *The combination of the academic achievement tests into two batteries for the prediction of academic success.*—The inter-correlations between the academic achievement tests for the total group of testees used in the investigation were computed. These are shown in Table 5.

TABLE 5

INTER-CORRELATIONS BETWEEN ACADEMIC ACHIEVEMENT TESTS

	Afrikaans	English	Mathematics	Physics	Chemistry
Afrikaans ..	—	·343	·346	·459	·294
English	—	—	·288	·264	·185
Mathematics	—	—	—	·643	·485
Physics	—	—	—	—	·586
Chemistry ..	—	—	—	—	—

Next, the inter-correlations in Table 5 and the coefficients of validity in Table 2 were used to compute multiple correlations for each course of study separately. It was observed that the inter-correlations for the B.A. and B.Sc. groups did not differ significantly, and therefore, the same figures could be used for both calculations.

Table 6 shows the multiple correlations and Beta weights computed by Aitken's method of Pivotal Condensation.[7]

TABLE 6

MULTIPLE CORRELATIONS AND BETA WEIGHTS FOR ACADEMIC ACHIEVEMENT TESTS AND B.A. (COMPOUND CRITERION).

Academic Achievement Tests	Five-test combination (1+2+3+4+5) Beta weights	Four-test combination (1+2+3+4) Beta weights	Three-test combination (1+2+3) Beta weights	Two-test combination (1+2) Beta weights
1. English	·331741	·326277	·333662	·343952
2. Chemistry ..	·220852	·213358	·214202	·268370
3. Physics	·109562	·085567	·097391	
4. Afrikaans ..	·032360	·031085		
5. Mathematics..	—·047300			
r mult.	·487	·485	·485	·478

From the results in Table 6 the suggestion can be made that, for all practical purposes, the Mathematics test (No. 5) and possibly also the Afrikaans test (No. 4) can be omitted, seeing that a five, four and three-test combination yield practically the same multiple *r*.

Table 7 shows the multiple correlations and Beta weights for B.Sc.

TABLE 7

MULTIPLE CORRELATIONS AND BETA WEIGHTS FOR ACADEMIC ACHIEVEMENT TESTS AND B.Sc. (COMPOUND CRITERION).

Academic Achievement Tests	Five-test combination (1+2+3+4+5) Beta weights	Four-test combination (1+2+3+4) Beta weights	Three-test combination (1+2+3) Beta weights	Two-test combination (1+2) Beta weights
1. Mathematics..	·272789	·293126	·301230	·335753
2. Physics	·103414	·106011	·144643	·232111
3. Chemistry ..	·183948	·184719	·187142	
4. Afrikaans ..	·067658	·098613		
5. English	·114898			
r mult.	·555	·545	·538	·517

The multiple correlations for the B.Sc. battery are considerably higher than those for the B.A. battery, as would be expected from the higher validities of the separate tests.

Further, the suggestion is made that a five-test combination or four-test combination in the order indicated on the table be used, for the prediction of academic progress in a B.Sc. direction of study.

Table 8 shows the multiple correlations for overseas tests and for two test batteries used by the afore-mentioned Student Guidance Committee.

It may be seen that the five-test combination for B.Sc. in this research compares favourably with, and even surpasses, the five-test combination of the University of Minnesota, and the four-test combination of the George

TABLE 8

MULTIPLE CORRELATIONS IN OTHER INVESTIGATIONS.

Investigation	Test combination	r mult.
George Peabody College for teachers[2]	seven	(·59) (·62)
	four	·57
University of Minnesota Dentistry School[6]	five	·54
Intelligence Tests :		
B.A.	two	·406
B.Sc.	two	·325
Bureau Tests :		
B.A.	five	·468

Peabody College for Teachers. The five-and four-test combinations of the B.A. battery are slightly lower than those of the overseas universities, but also compare fairly favourably.

It is also clear from Tables 6, 7 and 8 that even a two-test combination of academic achievement tests for B.A. and B.Sc. ensures a better prediction for academic success than intelligence tests, especially for B.Sc. A five-test combination of academic achievement tests also shows a better prediction of success than a five-test combination of Bureau tests for B.A.

III.—REFERENCES.

[1] ADKINS, D. C. (1947). *Construction and Analysis of Achievement Tests.* Washington D.C. : U.S. Government Printing Office.

[2] ANDERSON, S. B. (1953). Prediction and practice tests at the college level. *J. Appl. Psychol.*, 37, 256-259.

[3] BUREAU OF EDUCATIONAL RESEARCH AND SERVICE, EXTENSION DIVISION (1941). *Iowa Placement Examinations, Examiner's Manual.*

[4] EDUCATIONAL TESTING SERVICE (1955). *Co-operative School and College Ability Tests. Examiner's Manual.*

[5] GOUWS, D. J. (1957). *Die Akademiese aanpassing en vordering van eerstejaarstudente—'n Statisties-kliniese studie.* Unpublished D.Phil. thesis, University of Pretoria.

[6] LAYTON, W. L. (1953). Predicting success in dental school. *J. Appl. Psychol.*, 37, 251-255.

[7] MARITZ, J. S. (1949). The solution of normal equations by A. C. Aitken's method of pivotal condensation, with an introductory note on matrix theory. *Bull. Nat, Inst. Personnel Res.*, 5, 32-37.

[8] MOSIER, C. I., and MCQUITTY, J. V. (1940). Methods of item validation and abacs for item-test correlation and critical ratio of upper-lower difference. *Psychometrika*, 5, 57-65.

[9] NEL, B. F., and GOUWS, D. J. (1956). Some aspects of maladjustment of first-year university students in South Africa. *Int. Rev. Educ.*, 2, 174-188.

[10] PEARSON, R. (1949). Testing the Board's tests. *College Board Rev.*, No. 9, 111-112.

[11] RÜMKE, A. C. (1956). Statistische verwerking van studieresultaten aan de Rijksuniversiteit te Utrecht. *Paedagogische Studiën*, 188-189.

[12] STUDENTS GUIDANCE COMMITTEE (1957). *'n Ondersoek na die Akadamiese aanpassing en vordering van eerstejaarstudente aan die Universiteit van Pretoria.* Unpublished Report, University of Pretoria.

[13] WEBB, W. B. (1948-9). Occupational indecision among college students. *Occupations*, 27, 331-332.

[14] WENBERG, S. J. (1950). *The Student Veteran at Minnesota—Trends in Student Personnel Work.* Ed. E. G. Williamson. Minneapolis : University of Minnesota Press.

(*Manuscript received, May 21st,* 1958.)

A FACTORIAL STUDY OF SPEED, POWER AND RELATED TEMPERAMENT VARIABLES *

By G. L. MANGAN

(*Victoria University, Wellington, New Zealand.*)

SUMMARY. The fields of power, speed, persistence, accuracy and carefulness were investigated by factorizing thirty-eight sets of test scores among 200 13 and 14-year-old Ss. Rotation yielded five clearly identifiable and two less definite factors—content factors g + v and N, speed factors P, verbal tempo and fluency, and temperament factors persistence and speed-accuracy.

Persistence and speed-accuracy were identified as work-attitude or method-of-approach factors, which are closely linked with the difficulty of test material. Probably the operation of g and the content factors V, K and N are affected considerably by these work-habit components. The other speed factors isolated appear to be relatively unimportant.

I.—INTRODUCTION.

MOST comprehensive intelligence tests of the Binet type, as measures of power, call into operation such relatively independent aspects of problem-solving as level, various more specialized verbal, number and spatial abilities, and such temperament factors as speed, accuracy, persistence and carefulness, in underdetermined proportions.

These temperament or method-of-approach factors can be sampled without reference to the material differences of tests. Although it has been generally assumed that tests tend to cluster more according to their material content (V, N, etc.) than to their formal character, the fact that well-verified factors often have loadings on tests not composed of the usual materials suggests the operation of processes more general than subject matter classification.[39]

Factorial evidence about the nature and function of these factors is limited in some cases, in others difficult to interpret. Relevant data may be conveniently summarized as follows:

Speed.—Speed is defined, in Cattell's[6] terms, as 'rate of repetitive performance, where all the content material is perceptually given, through all cognitive levels.'

There is no clear evidence of a speed factor at a difficult level of cognition. While most initial research failed to extract a factor of speed as distinct from level [3, 14, 24, 33], the contrary finding, that speed and level are opposing modes of reaction, each of which has a tendency to prevail under certain conditions, has been often reported[6, 18, 19]. There are a number of speed factors, however, which are clearly recognizable: (1) at a low level of cognition [6, 17], for example in Tinker's[38] motor habit speed cluster, and in Thurstone's[35] 'speed of judgment' factor; (2) a reaction time speed cluster[19, 20]; (3) verbal fluency, or fluency of association[6, 15].

Tempo.—Tempo is defined as natural rate or rhythm, or basal speed when no urgent incentive is operating. It has been disclosed in two strongly correlated clusters centering round writing and tapping.[2] There is some evidence that there is a personal tempo and characteristic rate of movement[16, 32]; other data, however, have been interpreted as evidencing no psychomotor tempo of the individual, only agreements and affinities between comparable movements of the body[5, 21]. Rimoldi's[30] large-scale repetition and extension of Allport and Vernon's enquiry[2] is perhaps the most formative study ; although he analyzed his results into a number of multiple factors, all his inter-correlations were positive, and it seems justifiable to recognize a small general factor, together with group factors at special levels of activity, corresponding roughly with speed factors.

---

* This paper is condensed from the author's Ph.D. thesis, 1954. Thanks are due to Professor P. E. Vernon and A. Rodger of the University of London.

Persistence.—Persistence is defined as ' the ability to endure and continue to strive, for the sake of achievement, in spite of fatigue and discouragement.' Although there is strong indication of a general persistence factor[1, 23, 25] in that some groups of tests have underlying common features, there is some evidence that the trait is largely specific[29]. It is probable that there is a general factor running through persistence tests at the ideational level, and a group factor at the motor level.

Carefulness.—There is little evidence at the moment to justify the acceptance of carefulness as a consistent, general trait. Although the factor has been identified on a number of occasions[11, 14, 17] there is little external evidence of its nature. It does appear, however, to be similar to the ' caution ' factor Brown[4] and Manson[26] noted in intelligence test performances.

Accuracy.—Accuracy is ' that characteristic of measurable behaviour which is indicated by relative freedom from error, an error being considered as any deviation from an objectively correct standard.' From the numerous attempts to inter-relate speed and accuracy[17, 32, 34], two broad trends emerge :

(1) In tests in which intelligence plays a prominent part, or speeded tests at a high level of cognition, speed scores show positive correlations with accuracy scores.

(2) On the other hand, the negative correlations often found between speed and accuracy scores at motor tasks indicate the existence of an accuracy factor at the motor level.

In this research, an attempt is made to answer the following questions :

(1) Can speed, level, persistence, accuracy and carefulness factors be separated out under ordinary experimental conditions?

(2) Is speed or level the more appropriate measure of the general factor? If speed is a distinctive factor, how does it affect the material factor content of tests? Is it general to a great variety of tasks, is it highly specific, or does it reduce to certain better established factors W, P, etc. ?

(3) What are the inter-relationships of the factors, and what is the relative contribution of speed to the variance of power, compared with that of persistence, level, accuracy and carefulness, under the specified conditions?

II.—METHOD.

Thirty-two tests, which are described in Table 1, were used to measure the eleven factors postulated—power, persistence, carefulness, accuracy, clerical, perceptual, and motor speed and fluency, and verbal, non-verbal and motor tempo. These tests gave forty-five variables, but this number was reduced to thirty-eight by carrying out a Spearman-type analysis on each group of tests defining a factor. Those tests in each group showing least saturation with their common factor were combined ; such composites reduced the size of the matrix without detracting from the strength of the battery.

The battery of tests was administered in eight sessions. The subjects were 200 13 and 14-year-olds from a London comprehensive secondary school. The ability range was much greater than that usually found in selective secondary schools, although slightly restricted in the upper ability levels. The distribution of Duplex Test I.Qs. was approximately normal.

TABLE 1

THE TEST BATTERY.

Test	References*	Content	TL in minutes		Scoring Formula
			Fore ex.	Test	
1. Duplex level	9	Power	—	UNL	R
2. Duplex total	9	Power	—	UNL	R
Analogies Subtest-	—	Clerical speed	—	T–N	D
3. Dale Reasoning	7	Power	3	40	R
4. Prob. Arith.	—	Power	2	25	R
5. Matrices	—	Power	3	30	R
6. No. Series	23	Power	2	UNL	R
		Persistence	—	UNL	T
		Non- v. tempo	—	5	D
7. Incomplete words	3	Power	2	UNL	R
		Persistence	—	UNL	T
		Verbal tempo	—	5	D
8. Word building	23	Power	2	UNL	R
		Persistence	—	UNL	T
		Verbal tempo	—	5	D
9. Synonyms	3	Power	2	UNL	R
		Carefulness	—	7½	E
		Verbal tmepo	—	7½	D
10. Number square	23	Persistence	3	UNL	T
11. Diff. reading	23	Persistence	3	UNL	T
12. Cancellation	—	Carefulness	1	T–N	E
		Clerical speed	1	T–N	D
13. 3-higher	37	Carefulness	2	T–N	E
		Clerical speed	2	T–N	D
14. Measurement	17	Accuracy	2	T–N	A
		Motor speed	2	T–N	D
15. Copying para. 1	31	Motor tempo	—	T–N	D
16. Copying para. 2	31	Motor speed	—	T–N	D
17. Copying para. 3	31	Accuracy	—	T–N	A
18. Dotting 1	—	Motor speed	¼	4 15 sec. series	D
19. Dotting 2	—	Accuracy	¼	4 15 sec. series	A
20. Checking	—	Clerical speed	2	T–N	D
		Accuracy	—	T–N	A
21. Analogies	—	Clerical speed	2	T–N	D
22. Classification	37	Clerical speed	1	T–N	D
23. Scattered Xs.	37	Perceptual speed	2	4	D
24. Identical pict.	37	Perceptual speed	2	4	D
25. Fig. grouping	37	Perceptual speed	2	4	D
26. Birds	31	Fluency	—	3	D
27. Given letters	37	Fluency	1	3	D
28. Word series	31	Fluency	—	3	D
29. Copying Hs. 1	—	Motor tempo	—	4 15 sec. series	D
30. Copying Hs. 2	—	Motor speed	—	4 15 sec. series	D
31. Substitution	37	Non-v. tempo	2	5	D
32. Mental arithmetic	—	Non-v. tempo	—	T–N	D

TL = Time limit.
UNL = Unlimited time.
T–N = Subject recorded own time to complete the test.
R = Number of corre ct responses.
D = Distance covered in time limit, or time taken to cover material (T–N).
T = Total time taken.
A = Scored for accuracy.
E = Error scored.

* Many of the tests—Analogies for example—are well known. No references are given for these tests.

Pc

Pe

Ca

A

Cl
s

P

Fl

M

V

N

M

The testing covered a period of eight weeks. Normal class groupings of approximately thirty were retained, and the groups were rotated through the testing so that frequent rests were possible. No test session was longer than the usual class period of 60 minutes, and the tests were given in the same order to all subjects.

Scoring procedures varied. Three of the power tests were administered under ample time-limit conditions: the remaining five were untimed. In addition, three sections of the Duplex test were combined into a level score (cf. [12]). Power tests were scored for number and quality of responses. Persistence was scored as the time S was willing to spend on a task in hand when more attractive substitutes were available. Clerical speed tests, error scored (number right minus number wrong) provided carefulness scores. The accuracy tests were constructed in such a way that speed or accuracy preference was measured under different conditions of time and test content. Speed and accuracy aspects of work were equally stressed in the test instructions.

Timing of individual items of the speed and tempo tests was considered desirable since there is some evidence that wrong answers involve factors quite different from those in right answers[14] and any measure based on other than correct responses may introduce extraneous factors. Rate, therefore, was assessed as the amount of work completed in a given time, or by the time required to cover a certain amount of material, when the material was sufficiently easy for there to be practically no errors.

III.—RESULTS.

Analysis of the Data.—Factors were extracted from the 38×38 correlation matrix by Thurstone's centroid method. The highest correlation coefficient in each column was used as communality estimate. Extractions were continued until the product of the two highest centroid loadings (·08) was not appreciably greater than the standard error of the original correlation between the two variables (·06). The correlation matrix and the raw loadings are shown in Tables 2 and 3.

The most noticeable feature of the unrotated loadings, on inspection, was the dichotomous classification into speed and power elements. Any attempt to isolate all positive factors was, therefore, hardly feasible. However, speed, persistence, motor speed and tempo could be regarded as independent factors, since they involve little cognitive ability and are roughly orthogonal. The rotations thus aimed at isolating these, and then sub-classifying the other minor group factors, which, however, were not expected to be clear-cut. Orthogonal rotation procedures were followed, though as large a g as possible was retained. Direct graphical rotation necessitated eighteen rotations to stabilize the factors in the matrix. The rotated factor loadings are presented in Table 4.

TABLE 3

THE UNROTATED FACTOR MATRIX.

Test	Factor Loadings							Communality	
	I	II	III	IV	V	VI	VII	H_2est.	H_2obt.
1	475	—398	265	176	188	—091	161	714	555
2	639	—370	120	219	111	—148	314	714	740
3	509	—416	—110	—080	193	—038	154	568	513
4	572	—072	—226	—282	—105	143	086	432	502
5	543	—207	222	—102	—071	—137	—119	404	435
6	510	—311	—212	—142	068	232	—079	499	487
7	499	—416	038	094	151	096	—096	526	474
8	545	—036	254	197	—365	—207	—272	612	652
9	066	—318	146	—390	—346	089	—106	403	418
10	179	—543	243	—314	—123	073	—099	403	515
11	137	—343	200	—210	—189	050	086	376	266
12	—192	—366	253	—169	220	433	051	662	502
13	137	—278	—230	027	062	040	—121	254	177
14	173	—299	—169	148	088	117	—058	241	194
15	182	—101	—324	204	277	125	252	358	345
16	207	—334	—062	—130	—115	—044	—114	275	204
17	250	—128	—228	268	—062	—123	173	358	252
18	206	—291	—274	146	—178	221	066	348	308
19	496	317	105	140	060	095	—157	413	414
20	486	427	070	—101	—043	076	061	428	445
21	559	210	—109	—151	—047	272	—102	402	478
22	357	197	—171	077	194	075	—115	338	258
23	403	311	086	103	—176	295	235	434	450
24	134	520	113	082	—164	307	144	434	450
25	454	064	161	212	—108	107	112	410	317
26	422	—019	135	200	175	—030	—071	376	274
27	530	142	170	192	242	190	—131	367	481
28	455	198	062	—239	395	—269	—157	496	560
29	117	231	133	028	262	060	—132	223	182
30	307	333	082	—197	139	127	—092	345	295
31	436	—120	251	292	—169	—077	—207	556	430
32	291	267	086	181	—192	048	—076	277	241
33	469	244	—102	276	—247	—412	—185	662	630
34	233	276	—055	—197	—052	049	128	285	194
35	333	327	—249	161	—084	036	069	338	318
36	345	366	—348	—127	—203	175	—091	387	470
37	382	184	147	—249	305	—302	112	568	460
38	629	075	112	—206	—088	—222	094	568	523
Variance per cent.	15·91	8·65	3·39	3·72	3·53	3·28	2·02		

TABLE 4

THE ROTATED FACTOR MATRIX.

	Factor Loadings							Communality	
Test	I	II	III	IV	V	VI	VII	H_2rot.	H_2unrot
1	648	288	—087	—104	—073	—097	143	536	555
2	804	265	053	—119	—061	—003	—027	738	740
3	553	322	—061	—080	260	—066	—154	516	513
4	265	360	336	112	353	051	—221	503	502
5	383	459	019	181	032	158	139	436	435
6	306	365	128	—091	483	—010	—022	484	487
7	481	336	—037	—150	264	—003	191	474	474
8	316	342	182	131	—137	508	328	652	652
9	—201	610	002	—024	—006	041	—060	419	418
10	051	681	—170	—101	038	—080	046	515	515
11	056	485	—009	—093	—098	—067	—055	264	266
12	—410	—143	283	162	242	—388	—039	506	502
13	130	107	—105	—209	294	094	—024	179	176
14	216	081	—044	—273	250	030	053	198	194
15	358	—212	078	—237	257	—142	—160	348	345
16	108	360	—095	—086	150	148	—053	206	204
17	360	—065	111	—206	028	214	—146	256	252
18	146	153	194	—400	214	116	088	309	308
19	282	—036	343	289	139	100	320	414	414
20	204	023	472	420	063	019	050	448	445
21	171	179	462	240	371	041	079	478	478
22	242	—136	179	168	324	056	107	257	258
23	198	009	628	078	—053	—062	073	451	450
24	—029	—170	625	—091	—088	—019	133	455	450
25	361	106	366	007	—075	053	177	315	317
26	432	036	058	090	074	038	259	273	274
27	418	005	249	175	200	—073	410	481	481
28	363	032	—117	595	241	—013	058	562	560
29	082	—139	030	249	094	—126	248	176	182
30	063	022	243	410	185	—112	127	294	295
31	346	233	106	—019	—090	303	381	430	430
32	115	—026	354	128	—054	210	207	245	241
33	340	—081	175	245	—039	644	080	637	630
34	040	025	300	271	066	—032	—147	193	194
35	194	—205	401	106	144	201	—059	316	318
36	—047	—028	489	212	347	227	—109	471	470
37	376	—048	—058	541	006	—110	—098	461	460
38	423	325	211	398	—006	150	—101	521	523
	gv.	persistence	P	speed - accuracy	N	v.temp.	fl.		
Variance per cent.	10·8	6·8	7·0	5·8	3·9	3·4	2·9		

Description of the Factors.

FACTOR 1: g+v : ed.

Test No.	Test Name	Content	Loading
2	Duplex total	Power	·80
1	Duplex level	Power	·65
3	Dale	Power	·55
7	Matrices	Power	·48
26	Given letters	Fluency	·43
27	Word Series	Fluency	·42
38	Copying paragraphs	Motor tempo	·42
5	Incomplete words	Power	·38
37	Copying Hs.	Motor tempo	·38
15	Measurement	Accuracy	·36
17	Dotting	Accuracy	·36
25	Birds	Fluency	·36

The factor appears in most types of test and cuts across most of the scoring categories. It shows heaviest loadings on power, fluency, accuracy and certain motor tempo tests, though it enters into most of the verbal tests. It is principally involved, therefore, in reasoning and verbal tests, and is best described as a general power factor which includes part of the variance of fluency, tempo and accuracy tests. The general factor probably has a large verbal component.

FACTOR 2—Persistence.

Test No.	Test Name	Content	Loading
10	Number square+number series	Persistence	·68
9	Word building+incomplete words	Persistence	·61
11	Difficult reading	Persistence	·49
5	Incomplete words	Power	46
6	Number series	Power	·37
4	Problem arithmetic	Power	·36
8	Synonyms+Word building	Power	·34
7	Matrices	Power	·34

It is apparent that the factor involves time spent on tiresome asks, and is similar to Alexander's X factor, a temperament or attitude-to-work factor. It seems to affect results, however, only when the material is difficult, as indicated by the fact that it enters tests concerned with success at a difficult level of cognition, and is absent from tests dealing with easy material, especially when they are highly speeded.

FACTOR 3: Perceptual Speed.

Test No.	Test Name	Content	Loading
23	Identical pictures	Perceptual speed	·63
24	Figure grouping	Perceptual speed	·63
36	Mental arithmetic	Non-verbal tempo	·49
20	Classification+cancellation	Clerical speed	·47
21	3-higher+checking	Clerical speed	·46
35	Number series	Non-verbal tempo	·40
25	Birds	Fluency	·37

The factor is displayed through the perceptual, non-verbal, verbal and clerical speed tests involving matching or checking of easy material administered at speed. Thurstone[36] found that identical pictures, incomplete words, scattered Xs and figure grouping best defined P—'facility in perceiving details that are embedded in irrelevant material.' These results, therefore, tie in well with Thurstone's findings.

FACTOR 4: SPEED-ACCURACY.

Test No.	Test Name	Content	Loading
28	Copying Hs + copying para. 2	Motor speed	·60
37	Copying Hs 1	Motor tempo	·54
20	Classification + cancellation	Clerical speed	·42
30	Measurement	Motor speed	·41
13	Cancellation	Carefulness	—·21
15	Measurement	Accuracy	—·24
14	3-higher	Carefulness	—·27
18	Checking	Accuracy	—·40

It proved impossible to isolate accuracy as a factor orthogonal to speed, hence it seemed preferable to recognize a bipolar factor contrasting motor speed and tempo with accuracy. The eleven positive projections are limited to speed and tempo tests, mainly motor and clerical; the five significant negative loadings define a carefulness-accuracy sub-factor. There are no significant loadings on simple verbal or non-verbal cognitive or fluency tests. Most of the g : v power tests have intermediate positive or negative loadings, some bordering significance on the factor.

FACTOR 5: NUMERICAL.

Test No.	Test Name	Content	Loading
6	Number series	Power	·48
21	3-higher- + checking	Clerical speed	·37
36	Mental arithmetic	Non-verbal tempo	·35
4	Problem arithmetic	Power	·35
22	Scattered Xs.	Perceptual speed	·32

The factor is most clearly present in number series, 3-higher and mental and problem arithmetic, and loads all the numerical tests except number series + number square. N cuts across power and speed-accuracy, and has little verbal content; possibly it is slightly infused with the R or reasoning factor which many American researchers find in problem arithmetic and tests of the Matrices and Dale type.[13]

FACTOR 6: VERBAL TEMPO.

Test No.	Test Name	Content	Loading
33	Synonyms	Verbal tempo	·64
8	Synonyms + word building	Power	·51
31	Word building	Verbal tempo	·30
36	Mental Arithmetic	Non-verbal tempo	·23
32	Incomplete words	Verbal tempo	·21
17	Dotting 2	Accuracy	·21

The test list is not sufficiently extensive, and the variance is too slight to allow identification of the factor with any degree of confidence. It is suggestive, however, of a verbal tempo sub-factor.

FACTOR 7: FLUENCY.

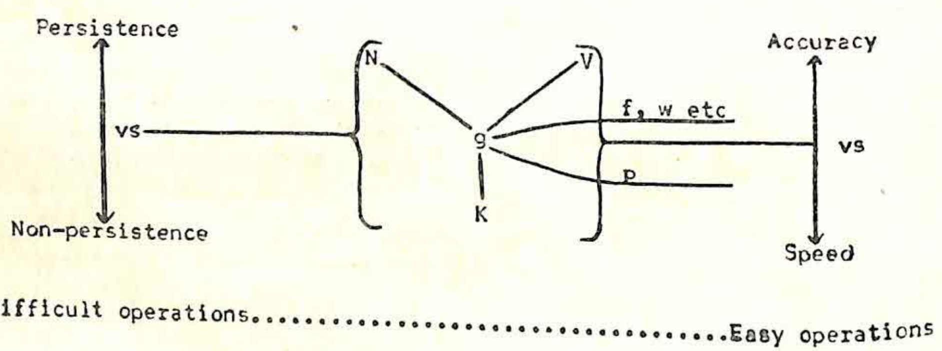
Test No.	Test No.	Content	Loading
27	Word Series	Fluency	·41
31	Word building	Verbal tempo	·38
8	Synonyms + word building ..	Power	·33
19	Duplex + analogies	Clerical speed	·32

It is very doubtful if a significant factor is represented. The factor is obscure and indeterminate, and though relatively unstable, is tentatively identified as a fluency sub-factor.

IV.—DISCUSSION.

In reviewing the different factors identified—g + v : ed, P, N, persistence and speed-accuracy, and two less definite factors, minor sub-groupings of the verbal tempo and fluency tests—it is clear that they can be fitted into a consistent, though incomplete, hierarchical picture of mental structure. The general factor isolated has been identified many times and has much in common with the g and v : ed factors. In this study, however, there appears no clear distinction as to which is the more representative g or gv—that measured by timed or untimed power tests. The untimed tests appear to have slightly higher saturations, but this is probably due to the fact that they are more lengthy. Untimed Duplex has by far the highest loading (·80); Dale and Matrices, however, both timed tests, have higher loadings (·55, ·48). than the remaining untimed tests. Speed appears to be of minor importance at a difficult cognitive level.

It is probable that g, with the other content factors, V, N and K, cannot be measured in isolation, but range between, or are conditioned by two work-attitude factors, persistence and speed-accuracy. This suggests a rather novel conception of ability factors (Figure 1). Persistence, which appears to have a great deal in common with Alexander's X factor, seems to affect results only when the material is difficult, and is non-significant in, or absent from tests incorporating easy material, particularly when they are highly speeded. Speed-accuracy, on the other hand, affects results only when easy cognitive or unskilled motor tasks are performed. Thus, when power tests are administered with ample or unlimited


Persistence
vs
Non-persistence
N
V
g
K
f, w etc
p
Accuracy
vs
Speed
Difficult operations..................................Easy operations

time, the ability measures will depend considerably on a temperamental or work-habit trait of persistence; if highly speeded tests are used, however another general attitude factor, speed-accuracy, intervenes. Further, even when unlimited time is allowed, the work attitude factor will change from speed-accuracy to persistence as the test progresses through easy to difficult material.

These work-attitude components have not revealed themselves in most previous research, except to some extent in the work of Davidson and Carroll,[8] Porebski[28] and Furneaux[12], since tests from which content factors have been derived have been graded usually at a similar level of speed and difficulty—that is, the tests have been largely confined to a single time or difficulty level, where work-attitude has been a constant factor. This condition applies to all content factors, and cuts across them.

This conclusion, however, is not incompatible with many previous research findings. Alexander concluded that the persistence or X factor he isolated was probably a temperament or work-attitude factor, but linked it with educational achievement rather than with the difficulty of test material. The probability that persistence is an important temperamental factor in ability has been stated a number of times[10, 39], but seldom in this context. Nor has a speed-accuracy factor been envisaged at the other end of the attitude scale being similarly called into operation when the material is easy.

It is fairly clear that the other speed factors isolated, P and fluency, are relatively distinct from speed-accuracy. Speedy perceptual identification, P, does not involve doing anything with the material. This appears to be the significant difference between it and speed-accuracy. When these factors are allowed for, the fluency (w or f) factor sinks into insignificance except in homogeneous groups. As previously indicated by Rogers[31] and Morgan[27], quickness and richness of mental associations do not stand out as clear components.

V.—REFERENCES.

[1] Alexander, W. P. (1935). Intelligence, concrete and abstract. *Brit. J. Psychol. Monogr: Suppl.* 6, No. 19.

[2] Allport, G. W. and Vernon, P. E. (1933). *Studies in Expressive Movement.* New York. Macmillan.

[3] Bernstein, E. (1924). Quickness and intelligence. *Brit. J. Psychol., Monogr. Suppl.,* 3, No. 7.

[4] Brown, W. M. (1924). A study of the 'caution' factor and its importance in intelligence test performance. *Amer. J. Psychol.,* 35, 368-386.

[5] Campbell, M. (1934). The cognitive aspects of motor performances and their bearing on general motor ability. *J. Exper. Psychol.,* 19, 323-333.

[6] Cattell, R. B. (1943). The measurment of adult intelligence. *Psychol. Bull.,* 40, 153-193.

[7] Dale, A. B. (1926). Group tests in reasoning ability. *Brit. J. Psychol.* 14.

[8] Davidson, W. M., and Carroll, J. B. (1945). Speed and level components in time-limit scores: a factor analysis. *Educ. Psychol. Measmt.,* 5, 411-428.

[9] Earle F. M. (1948). *The Duplex Series of Ability Tests for Age Groups* 10 *to* 14. London: Harrap.

[10] Eysenck, H. J. (1947). *Dimensions of Personality.* London: Kegan, Paul, French, Trubner.

[11] Fruchter, B. (1950). Error scores as a measure of carefulness. *J. Educ. Psychol.,* 41, 279-291.

[12] Furneaux, W. D. (1948). Some factors affecting the design of 'g' with particular reference to the relation of 'speed' and 'power.' *Proc. Twelfth Inter. Congr. Psychol.,* Edinburgh.

[13] Guilford, J. P. (1941). The difficulty of a test and its factor composition. *Psychometrika,* 6, 67-77.

[14] Guilford, J. P., and Lacey, J. I. (1947). Printed classification tests. *Army Air Forces Aviat. Prog. Res., Rep. No.* 5, Washington, D.C.: U.S. Govt. Printing Office.

[15] Hargreaves, H. L. (1927). The 'faculty' of imagination. *Brit. J. Psychol., Monogr. Suppl.*, 3, No. 10.

[16] Harrison, R. (1941). Personal tempo and the inter-relationships of voluntary and maximal rates of movement. *J. Gen. Psychol.*, 24, 343-379.

[17] Himmelweit, H. (1946). Speed and accuracy of work as related to temperament. *Brit. J. Psychol.*, 36, 132-144.

[18] Kelley, T. L. (1927). *Interpretation of Educational Measurements.* New York : World Book Co., Yonkers-on-Hudson.

[19] Kennedy, M. (1930). Speed as a personality trait. *J. Soc. Psychol.*, 1, 286-299.

[20] Lanier, L. H. (1934). The inter-relations of speed of reaction measurements. *J. Exper. Psychol.*, 17, 371-399.

[21] Lewitan, C. (1927). Untersuchungen über das Allgemeine Psycho-motorische Tempo. *Ztsch. f. Pshycol.*, 101.

[22] Longstaff, H. P., and Porter, J. P. (1928). Speed and accuracy as factors in objective tests in general psychology. *J. Applied Psychol.*, 12, 636-642.

[23] MacArthur, R. S. (1950). *An Experimental Investigation of Persistence and its Measurement at the Secondary School Level.* Ph.D. thesis, University of London.

[24] McFarland, R. A. (1928). The role of speed in mental ability. *Psychol. Bull.*, 25, 595-612.

[25] Maller, J. B. (1934). General and specific factors in character. *J. Soc, Psychol.*, 5, 97-102.

[26] Manson, G. E. (1925). Personality differences in intelligence test performances. *J. Appl. Psychol.*, 9, 230-235.

[27] Morgan, G. A. (1953). *A Study of Verbal Abilities.* Ph.D. thesis, University of London.

[28] Porebski, O. (1952). *A Psychological and Statistical Study of Speed and Power as Variables of Human Ability.* M.A. thesis, University of London.

[29] Porter, J. P. (1933). A comparative study of some measures of persistence. *Psychol. Bull.*, 30.

[30] Rimoldi, H. J. A. (1951). Personal tempo. *J. Abn. Soc. Psychol.*, 46, 283-303.

[31] Rogers, C. A. (1952). *A Factorial Study of Verbal Fluency and Related Dimensions of Personality.* Ph.D. thesis, University of London.

[32] Slater, P. (1938). Speed of work in intelligence tests. *Brit. J. Psychol.*, 29, 55-68.

[33] Spearman, C. 1927). *The Abilities of Man.* London : Macmillan.

[34] Thorndike, E. L. (1915). The relation between speed and accuracy in addition. *J. Educ. Psychol.*, 5, 537-541.

[35] Thurstone, L. L. (1935). *The Vectors of Mind.* Chicago : University of Chicago Press,

[36] Thurstone, L. L. (1947). *Multiple Factor Analysis.* Chicago : University of Chicago Press.

[37] Thurstone, L. L., and Thurstone, T. G. (1941). Factorial studies of intelligence. *Psychometric Monogr.* No. 2. University of Chicago Press.

[38] Tinker, M. A. (1931). The significance of speed in test response. *Psychol. Rev.*, 38, 450-454.

[39] Vernon, P. E. (1950). *The Structure of Human Abilities.* London : Methuen,.

(*Manuscript received, June* 17*th*, 1958.)

CATEGORIES IN RATING SCALES: HOW MANY AND HOW BIG?

BY FRANK SANDON

(*Chief Examiner, Somerset Education Committee.*)

SUMMARY. Various bases have been proposed to give the number and relative frequencies of the categories used in rating scales. There is always the possibility of error of estimate and it is misleading to give ratings that may be grossly at fault.

From a consideration of:

(*a*) the possibility of a rater being able to estimate fairly accurately the typical performance of a ratee,

(*b*) the possibility of the rater being able to estimate fairly accurately how that typical performance of the ratee compares with that of all ratees, and

(*c*) the purpose of the rating scales,

suggestions are made to show how rules can be derived to indicate how many categories (literal grades) should be used and what should be the relative frequencies in these.

The rules depend on the assumption that there is a normal distribution of the quality reviewed and on the requirement that any literal grade that may be assigned to any individual should not be more than one grade out in 95 per cent. of the cases. They can be modified for finer discrimination, or if we allow for other probabilities or for wider errors; and it is suggested that the conclusions are valid when the assumption of normality does not hold.

The more accurately an estimate can predict the criterion, the finer can be the classification.

I.—INTRODUCTION.

TEACHERS are asked from time to time to give estimates about their pupils. In some cases the estimates are to be made in the form of ranked lists. But very often the teacher finds difficulty in the fine discrimination necessary for preparing such rankings, and he is invited, as often in compiling School Record Cards, to use a Rating Scale. In the G.C.E. he may be asked to estimate the pupils' probable performance on a 21-point rating scale, using the 'grades' 0, 5, 10, . . . 95, 100. In other cases he may be asked to use a five-point rating literal scale, a device first proposed by Burt: he suggested (Burt, 1917) the system Very Good, Very Fair, Medium, Very Poor, Very Bad, with relative percentage frequencies of 7 : 24 : 38 : 24 : 7. Yates and Pidgeon (1957) have pointed out that a three-point rating scale seems for some purposes to be elaborate enough.

Recently, the writer had occasion to prepare estimates for the G.C.E. for a very weak class of pupils, few of whom were expected to pass. In various preliminary tests the relative performance of these pupils were much interchanged, as one or another happened to hit on a method that earned some marks. It was found very difficult to assign any reasonable estimate of the prospects of each pupil.

A consideration of the above features suggested that:

(*a*) the formal problem of the assignment of the rating scale is the same whether it has to be given in marks or on a literal (or verbal) scale;

(*b*) the accuracy of our knowledge of the pupil will suggest how many categories should be available to use in the scale;

(*c*) no assignment to a rating scale can be expected to be absolutely correct.

These ideas are developed below.

II.—ASSUMPTIONS.

We shall, following other workers, talk of a rating scale as divided into categories or grades, and consider that the measure of the attribute to be rated follows a normal distribution among the ratees. Thus Burt's five categories are those bounded by values lying at positions of approximately $-1{\cdot}5$, $-0{\cdot}5$, $+0{\cdot}5$, and $+1{\cdot}5$ in standard deviation units. The National Foundation for Educational Research suggested (see, e.g., Hamley), that the five categories of its rating scales should have percentage frequencies of 5 : 25 : 40 : 25: 5. These are blocks of the normal curve of approximate widths 1·1 times the S.D. In various other cases (e.g., Vernon, 1938 ; Vernon and Parry, 1949 ; and Sandon, 1952) the instructions for the use of rating scales include suggested relative percentage frequencies (rounded to integral values) of the numbers to be assigned to each of the grades. These frequencies of the categories have been computed as blocks of the normal curve all of equal width. It is found that the boundaries are practically equally spaced even if, instead of making this our explicit objective, we try to get the positions of the means of the various categories equally spaced (see Appendix A). In the following we shall, while still taking our categories as of equal width, adopt a new principle, that of assigning them in such a way that, for a specified misplacement in rating, the probability of error does not exceed a certain pre-assigned value. We shall consider two types of case, the 'even' cases, where there are equal numbers of blocks on either side of the mean of the normal distribution, and the 'odd' cases, where the middle block is of the same width as the others, but straddles the mean symmetrically.

III.—ASSIGNING THE CATEGORY.

Let us consider how we rate any particular pupil, Smith. We have observed the attribute in Smith on a number of occasions. He does not always show the same amount of the attribute but we think that we can place him on the rating scale with reasonable accuracy. There are various categories of the scale in which he might be placed, though some of the available assignments would be quite inappropriate. If we define what we mean by making an error in our assignment we can determine the degree of probability that our rating will be in error. The procedure is indicated in Appendix B. If we calculate, using the method there indicated, the probability that, e.g., we do not make any error greater than that of putting Smith more than one category out, we reach the results shown in Table 1. Adopting the usual convention of statisticians and

TABLE 1

PROBABILITY OF ERROR FOR CATEGORIES OF VARIOUS WIDTHS.

Width	'Even' Cases		"Odd' Cases	
	Probability	Relative Frequency of Categories	Probability	Relative Frequency of Categories
1σ	·299	2 : 14 : 34 : 34 : 14 : 2	·299	7 : 24 : 38 : 24 : 7
$1{\cdot}2\sigma$	·216	1 : 11 : 38 : 38 : 11 : 1	·217	3 : 24 : 46 : 24 : 3
$1{\cdot}4\sigma$	·153	8 : 42 : 42 : 8	·154	2 : 22 : 52 : 22 : 2
$1{\cdot}6\sigma$	·105	5 : 45 : 45 : 5	·109	1 : 20 : 58 : 20 : 1
$1{\cdot}8\sigma$	·070	4 : 46 : 46 : 4	·077	$18\frac{1}{2}$: 63 : $18\frac{1}{2}$
$2{\cdot}0\sigma$	·049	2 : 48 : 48 : 2	·054	16 : 68 : 16

N.B.—The relative frequencies have been rounded.

taking a probability of ·05 (1 in 20, or 19 to 1 against) as the limit of significance, we see that the sets of categories appropriate to this chance of error of ·05 are approximately those given in the last line of the table.

IV.—EFFECT OF CORRELATION WITH A CRITERION.

We have based the above on the assumption that we are assigning Smith on our rating scale with reasonable accuracy. But, as already suggested, the accuracy of our estimates varies. If we consider the correlation of our estimate with a criterion, then we shall appreciate that we can make more reasonable estimates when the correlation is higher. We note that the fewer categories we have the smaller becomes the probability of an error of ' at least one category out.' This is, perhaps, obvious ; for in the limiting case of two categories only we cannot be more than one category out and the probability of such an error is thus zero. The categories corresponding to a probability of ·05 for not more than one category out are all concentrated about Smith's proper level on the criterion, so that if we can use the whole possible range of the criterion we can pack in more categories than those suggested for rating Smith only. The possible categories in such cases are shown in Table 2 ; their derivation is explained in Appendix C.

TABLE 2

TABLE TO SHOW RELATIVE FREQUENCIES IN CATEGORIES OF MARGINAL DISTRIBUTION FOR ARRAYS OF THE PREDICTOR OF SPECIFIED ERROR (P= ·05) CORRELATING AT VARIOUS LEVELS WITH THE CRITERION.

Value of r.	' Even ' Cases	' Odd ' Cases
0	2 : 48 : 48 : 2	16 : 68 : 16
·4	3 : 47 : 47 : 3	18 : 64 : 18
·6	5 : 45 : 45 : 5	21 : 58 : 21
·7	8 : 42 : 42 : 8	2 : 22 : 52 : 22 : 2
·8	1 : 11 : 38 : 38 : 11 : 1	3½ : 24 : 45 : 24 : 3½
·85	2 : 13 : 35 : 35 : 13 : 2	6 : 24 : 40 : 24 : 6
·9	4 : 15 : 31 : 31 : 15 : 4	1½ : 8 : 23½ : 34 : 23½ : 8 : 1½
·95	3 : 7 : 16 : 23 : 23 : 16 : 7 : 3	2 : 4 : 12 : 20 : 24 : 20 : 12 : 4 : 2

V.—USE OF RATING SCALES OF DIFFERENT SUB-DIVISIONS.

If we are considering the kind of estimate that we have to make in the G.C.E. where, although Smith is not steady in his performance in the school, we have a pretty fair idea of his true position in the criterion, then we have the case of a number ot categories available for our rating, of which a few only are appropriate for Smith : this is true even in the situation referred to at the beginning of this paper. This is the kind of problem for which the categories given in the last line of Table 2 are appropriate. But even here we do not get the 21-point scale prescribed by the G.C.E. Regulations. Also it must be remembered that, although the coefficient of correlation between the predictor and the criterion for the pupils of any one rater is probably nothing like as much as ·95, yet it is very seldom that he is dealing with pupils covering the whole range of ability. The greater homogeneity of his class necessarily results in smaller values for the correlation (Sandon, 1936). One of the difficulties in using a scale of this kind is that one cannot assume that one's pupils are a random sample of all those intended to be covered by the large number of categories.

The experienced G.C.E. teacher, with the numerical results of his classes of earlier years to go on, has some idea of where Smith and his form-mates should be placed. But for most rating scales the rater is not in a position to know how far or in what direction his sample is biased. The difficulty that the recipient has in using ratings submitted by different schools is the same problem as that of using ranking or any other teachers' estimates. Various solutions have been suggested for this (Sandon, 1956 ; Vernon, 1957) ; none of them is satisfactory.

In school terminal reports it is a common practice to give besides, or alternative to, a numerical 'percentage' mark or a class ranking, a literal grade, possibly with + or – attached. As the term-by-term correlation in any school subject is of the order of ·7 (e.g., Sandon, 1933), we can use Table 2 to indicate how, if we adopt the definition of error above, we can use some five literal grades that will give a fairly steady value for Smith in successive terms.

A five-point scale is often used in Teachers' Certificate Examinations, where A, for example, indicates an undoubted distinction and E a certain failure. Although percentage figures are suggested here yet, as a recent *Times Educational Supplement* correspondent has said, the system "implies a fine accuracy of assessment which is probably spurious." This comment illustrates once again the two difficulties we have to face—that we can only place Smith with some fuzziness at a point on our scale, and that we do not know how our scale compares with that of other assessors.

In many cases the proper rating of Smith is much vaguer than for a G.C.E. subject or for a terminal report. If we are asked to make a prognosis of Smith's success in a grammar school course, we are reminded by Yates and Pidgeon (1957), that a three-point scale is probably elaborate enough : they suggest using A for children eminently suitable, B for the possibly suitable, and C for the definitely unsuitable. To quote : "The three-point scale permits the use of the middle category for those children about whose chance of success in a grammar school they (the teachers) feel unable to predict with any certainty. Any extension of the number of categories beyond this is regarded, in our experience, with misgiving." In such a case, and in most cases of personality, or clinical (Jones, 1938), assessment, the rater has only a general idea of the numerical value of the measure behind the attribute and the correlation with any criterion is very low. And for cases of low correlation, Table 2 indicates that it is appropriate to use only a few categories in our rating scale.

VI.—REMOVING THE ASSUMPTIONS.

It must be remembered that, throughout the above, we have assumed a normal frequency surface. It has been pointed out (e.g., Burt, 1943 ; Sandon, 1952) that for many qualities, particularly those involving personality assessments, such an assumption is not justified. It may not be justified for either the predictor or for the criterion or for both, so that the suggestions about the relative frequencies of the categories, or even of the number of the categories, may not always be valid. But it will probably be found in practice that the general idea of a few categories for weakly correlated predictors and of more for highly correlated ones will give good service.

If we do not take as valid the assumption that the boundaries of the categories are derived from equally spaced abscissae of the normal surface, then we have merely an alternative form of the assumption that the distribution is non-normal.

Here, however, we should remark on a practical feature, that of the limited sample with which the rater is working. The class teacher as a rule is dealing with what statistically is a small sample—the number of individuals in his distribution is often of the order of thirty only. In such a case, the distribution will not exhibit exactly the features of a theoretical distribution. In particular, the distances (on the basic measure) between successive individuals, when ranked according to this measure, will not follow precisely those appropriate to any such distribution, but instead will be closely clustered at some places and will reveal wide gaps at others. The experienced rater will probably in such cases not attempt to follow the prescribed frequencies laid down as suggestions by the user, but will take them merely as a general guide and will use these clusters and gaps to obtain the required number of categories, and the experienced user will appreciate that the rater has found it convenient to draw his dividing lines where there seemed to be best evidence of real discrimination. Usually, of course, the clustering will be most intense near the mode. In consequence, for most practical situations, to ask for the use of an 'even' set of categories will mean asking the rater to draw dividing lines where it is least likely that any natural discrimination will reveal itself. As we have seen in Table 1, there is little to choose, from the point of view of probability of error, between the corresponding 'odd' and 'even' cases. For some purposes it may be necessary to ask for one of the 'even' categories, but, on the whole, it is found in practice that teachers, probably unknowingly for the reason just stated, prefer to work with one of the 'odd" set of categories.

VII.—APPENDICES.

APPENDIX A.

Kelley (1950) has said that it is impossible to have category means such that all differences of distance between adjacent means are equal. In practice this is not true. Consider an 'even' case and let us suppose that the middle of the first category on the positive side has its upper boundary at +1 (measured in terms of the S.D.). Then, if the means are to be equally spaced, the other boundaries are as shown in the following table :

TABLE 3

Boundary	Successive Blocks, Counting from Mean				
	1	2	3	4	5
Lower	0·0	1·0	2·0	3·0	3·7
Mean	0·46	1·38	2·30	3·22	4·14
Upper	1·0	2·0	3·0	3·7	—

The distance between successive means is ·92.

There is no category starting at 3·7 that can have its mean (i.e., its C.G.) at 4·14. But note that the first three blocks, accounting for all but about one case in a thousand, are of equal width, to the accuracy shown. Similar results are obtained for other widths of 'even' cases : the 'odd' cases likewise give impossible boundaries at the ends of their range. An alternative basis of determining categories has been considered by Kelley (1950) : once again the boundaries are effectively equally spaced.

APPENDIX B.

Suppose that we have a distribution, where successive categories A_1, A_2, A_3, . . . A_r, . . . have relative frequencies a_1, a_2, a_3, . . . , a_r, . . . which total 1·0. Then an individual correctly assigned to A_r may be assigned in error to any other category.

The chance that he is correctly assigned is a_r: the chance that he is not more than one category out is $a_{r-1}+a_r+a_{r+1}$. The total probability that the cases assigned by the rater are at least two categories out is, therefore, $\Sigma a_r(1-a_{r-1}-a_r-a_{r+1})$.

In Table 1 the percentage frequencies of the various categories of specified width have been, as usual, given as integers whenever possible, and the probabilities have been calculated for these rounded values. The exact probabilities for categories of the precise frequency for the widths stated are, however, as follows:

TABLE 4

Type of Case	Standard Category Width Quoted in Table 1					
	1σ	$1{\cdot}2\sigma$	$1{\cdot}4\sigma$	$1{\cdot}6\sigma$	$1{\cdot}8\sigma$	$2{\cdot}0\sigma$
Even	·2960	·2264	·1472	·0950	·0768	·0392
Odd	·2986	·2010	·1568	·1114	·0685	·0512

It will be noticed from Table 1 that the probability of error is always smaller for an 'even' case than for the precisely corresponding 'odd' case, though the difference is small. When we round the frequencies these systematic differences disappear.

The precise values of the widths corresponding to a P of ·05 are, by interpolation, $1{\cdot}949\sigma$ for the 'even' case, and $2{\cdot}038\sigma$ for the 'odd' case.

It will also be noted in Table 1 that our rounding rule does not give systematically more categories to one type of case than to the other.

In the practical application of any instructions about categories, the rater will be dealing with groups of individuals such that only very exceptionally will any suggested specified percentage frequencies give integral values for the number of individuals per category. Inevitably, therefore, there will be slight divergences from these suggested frequencies, with corresponding differences in the various probabilities. As pointed out at the end of Section VI the rater will regard the specified percentages as mere suggestions, and though he should attempt to get his relative proportions of numbers per category fairly close to the suggestions, he will be ill advised to attempt to force his data to the closest arithmetic fit of numbers.

There is a further point that we may consider. Suppose that we determine the probability of a wider error than that which we have just dealt with. As an illustration, suppose Category A_r be divided into three sub-categories of equal width, each one-third of A_r, and suppose that the frequencies of these be ${}_{-1}a_r$, ${}_0a_r$, ${}_{+1}a_r$, where ${}_{-1}a_r+{}_0a_r+{}_{+1}a_r=a_r$. Then the chance that a case is not more than one of the coarser categories out is $\Sigma({}_{-1}a_r+{}_0a_r+{}_{+1}a_r)\ (1-{}_{-1}a_{r-1}-\ .\ .\ .\ -{}_{+1}a_{r+1})$ and the chance that a case is not more than three out of the finer categories is $\Sigma_0a_r\ (1-{}_0a_{r-1}-\ .\ .\ .\ -{}_0a_{r+1})$. The second probability is larger by $\Sigma_0a_r\ ({}_{-1}a_{r-1}+{}_{+1}a_{r+1})$. This result is general. As a numerical example, consider the 'even' case where the original probability was ·05 for an error of not more than one category out. The original width is, as seen above, 1·949. Then the probability of an error of not more than two categories out when the width is 1·949/2 or ·97 can be computed to be ·09 and of an error of not more than three categories out when the width is 1·949/3 or ·65 to be ·11. We should, therefore, for the same total probability of error, have to consider, as acceptable, errors greater than these if we sub-divided the range to get more categories.

APPENDIX C.

The S.D. of the homoscedastic arrays in a correlation table of a normal surface of correlation is $S\sqrt{(1-r^2)}$ where S is the s.d. of the marginal distribution and r is the coefficient of correlation. But our blocks of Table 1 are of width $\alpha\sigma$ where α is the approppriate multiple (e.g., 1·0, 1·2, . . . 2.0) of σ, the s.d. of that array, so that the widths of the separate arrays that were used in calculating Table 1 are only $\alpha S\sqrt{(1-r^2)}$ wide when we derive them from the marginal frequency. This gives us the more marginal categories the closer the arrays cluster about the regression line. Table 2 is derived by calculating the corresponding frequencies, rounding again as before. It may be noted that the assumption of homoscedasticity forces us to use for our categories blocks of equal width alike for the separate arrays and for the marginal frequency distribution.

VIII.—REFERENCES.

BURT, C. (1917). *The Distribution and Relations of Educational Abilities.* London : P. S. King.

BURT, C. (1943). Ability and income. *Brit. J. Educ. Psychol.*, 13, 83-98.

HAMLEY, H. R., *et al.* (1937). *The Educational Guidance of the School Child.* London : Evans' Bros.

JONES, R. H. (1938). Physical indices and clinical assessments of the nutrition of school children. *J. Roy. Stat. Soc.*, 101, 1-23.

KELLEY, T. L. (1950). The use of literal grades. *J. Educ. Psychol.*, 41, 488-492.

SANDON, F. (1933). Progress through a secondary school as measured by school marks. *Brit. J. Educ. Psychol.*, 3, 267-290.

SANDON, F. (1936). Selection by a nearly perfect examination. *Annals of Eugenics*, 7, 65-85.

SANDON, F. (1952). The distribution of personality assessments. *Educ. Review*, 5, 53-65.

SANDON, F. (1956). A regression control chart for use in personnel selection. *Applied Stats.*, 5, 20-31.

VERNON, P. E. (1938). *The Assessment of Psychological Qualities by Verbal Methods.* London : H.M.S.O.

VERNON, P. E., and PARRY, J. B. (1949). *Personnel Selection in the British Forces.* London : University of London Press.

VERNON, P. E. (1957). *Secondary School Selection : A British Psychological Society Inquiry.* London : Methuen.

YATES, A., and PIDGEON, D. A. : NATIONAL FOUNDATION FOR EDUCATIONAL RESEARCH IN ENGLAND AND WALES (1957). *Admission to Grammar Schools.* London : Newnes' Educational Publishing Co.

(*Manuscript received, July 25th*, 1958.)

RESEARCH NOTE

The Importance of Sand in the World Technique: An Experiment.

By RUTH BOWYER
(*Pyschology Department, University of Bristol.*)

I.—Introduction.

One of the major differences between the Lowenfeld World technique and Dr. Charlotte Buhler's modification of it, is the place given to sand. Whereas Dr. Margaret Lowenfeld always uses a sand-filled tray in which the subject makes his construction, Dr. Buhler* advocates " at least six feet square of floor space, preferably in the centre of the room so that he can walk around his work."'"Adolescents and adults often prefer to work on a large table." It is true that Dr. Buhler sometimes provides a sand-box which children may use if they prefer, but she regards this as more interesting for play therapy than for diagnosis. It seems strange that her wide experience has not led her to give more consideration to standard diagnostic principles for the use of sand.

In the gradual collection of age norms at Bristol, where to date Worlds have been made by some 300 persons (apart from those obtained in play therapy) it has been found that sand is used destructively, expressively, symbolically, or constructively; is ignored except as a base; or is reacted against, to the extent of covering it over with cardboard. Between the ages of two and four years children tend to pour sand over people and things, or to bang the sand, sometimes with the words " Down, down ! " Burying also occurs, not always with hostile intent. Sometimes the vigour of the young child prevents him from balancing an animal on the surface of the sand, or perhaps non-recognition of the surface boundary causes him to thrust objects deep into the sand. Mrs. Eve Lewis, in a letter about Worlds from Exeter Child Guidance Clinic, writes : " I have often observed the careful burying of some toy to indicate, not aggression, but acceptance of that which the toy symbolized in the collective unconscious. The child usually asks that it be left for him to find again next time." After the age of 4 or 5 years, almost to the teens, there is relatively little use of sand except in clinic cases. From 13-15 years, or earlier, where the mental age is advanced, there begins to appear a constructive use of sand, for hills and valleys, roads and rivers, waves of the sea, or furrows in a ploughed field. This paper gives an account of an experiment designed to answer the question of the diagnostic import of such a constructive use of sand.

II.—The Hypotheses.

From experience of subjects who manipulated sand to form hills, valleys, roads and rivers, it seemed that this is a fairly sophisticated response, rarely occurring with children under eleven years. Younger children who used the sand in this way did so only after they had been coming for therapy for some time, with the exception of one 7-year-old boy with a Stanford-Binet I.Q. of 158, who used the sand constructively on his first visit.

The effect of the constructive use of sand is to create a more realistic, life-like product. Following Lewin's concept of differentiation as the hallmark of development, which was a useful frame of reference in ordering the normative data,† we can see that the moulding of the landscape to secure a three-dimensional quality gives increased life-space, variety and organisation. The sand base now forms an interdependent relationship with the toy objects, e.g., tractors, ships, trees. Intercommunication is secured by road, river or tunnel. From consideration of the concept of differentiation, as well as from observation of the behaviour of people who made Worlds, it seemed probable that further evidence would support the hypothesis that the constructive use of sand is a sign of good intelligence.

---

* Buhler, Charlotte. *Journal of Child Psychiatry*, 1951, 2, pp. 69-82.

† Bowyer, Ruth. *A Normative Study of Sand Tray Worlds.* Paper read at B.P.S. Annual General Meeting, Manchester, 1956.

A further hypothesis arises as to the quality of this intelligence. It has to be taken into account that a person who proceeds to move the sand in response to the instructions " make what you like, in here, with these toys " is going beyond what many people regard as given. He is using initiative, he is to some extent restructuring his world, he is making an even more imaginative response than that of the person who creates a scene with the toys on static sand. The further hypothesis, therefore, is that the constructive use of sand bears in part a similar interpretation to the Rorschach movement response (M). Both, i.e., sand construction and perception of human movement in the blot, are creative rather than passive responses. Both signify availability of inner resources of imagination and intellect.

There is no evidence that the constructive use of sand has also the social implications of the M responses, namely that there is sufficient interest in persons for social relatedness. Indeed, even in the Rorschach, the M responses needs support from the rest of the pattern for such a prediction. In the World technique there is no suggestion that the equivalence with M goes beyond a relation with creative intelligence. To the extent that imagination and intelligence are valuable in themselves, the constructive use of sand is a good prognostic sign, i.e. " other things being equal."

III.—THE EXPERIMENT.

The opportunity to test the hypothesis that the constructive use of sand is an intelligent response came when Dr. Lumsden Walker and his staff at Hortham Hospital, Gloucestershire, kindly allowed the World technique to be used with their mentally defective patients, in furtherance of developmental norms. 216 patients made Worlds and there was a control group of fifty intelligent subjects. The criterion of intelligence was high school or university education. The age range of the Hortham patients was 7–58 veras, the majority being 20–35 years old. The control subjects were from 14–50 years of age, the majority being 20–25 years old.

The Hortham patients made Worlds in the hospital Board Room, where it was not practicable to offer the free use of water. Therefore, the sand was kept at a uniform dampness, sufficient for easy moulding. The control group made Worlds in Bristol University Department of Psychology, and half of these control subjects (namely those who were also given the Rorschach test) were presented with the sand dampened as for the Hortham patients. The remaining control subjects had made their Worlds earlier in accordance with the usual instructions to use water if they wished.

The Rorschach test was administered to twenty-one of the control group, on the same day or within a week of making a World. The reason for stopping at that number was that no subject failed to give M responses, and so it seemed that the expectation of M from this group was sufficiently supported by the sample of twenty-one to justify the testing of an equal number of mentally defective subjects. Only high grade patients were chosen, namely one who used the sand constructively, two who did so tentatively, and any others sent by the hospital staff, provided the age range was the same as for the twenty-one control subjects. It was necessary to exclude children, since there could not be much expectation of M responses from them. The reason for giving the Rorschach test to high grade patients only was to secure Rorschach protocols with a sufficient number and range of responses to make the search for M meaningful. It also seemed best to put the hypothesis of few M responses from the Hortham group to as hard a test as possible.

IV.—RESULTS.

The results were as follows :

A.—*Numbers using sand constructively* :

Hortham Group (216).	*Control Group* (50)
1	48

The only Hortham patient who used the sand constructively was a woman with a Wechsler I.Q. of 95 who is a ' social risk ' case. Two other high grade patients made beginnings, with an inch of path poked out sketchily in the sand with a finger.

Of the two members of the control group who did not use sand at all, one was the youngest subject (age 14) and the other was a university student who remarked that from time to time in her childhood she attended a psychiatrist about skin troubles which assail her whenever she is upset.

B.— *Rorschach M Responses.*

Hortham Group (21)		Control Group (21)
Persons with M	6	21
Total No. of M in group	7	135
Average No. of M per person	0·33	6·4

Of the Hortham group, five persons had one M response each, and one man had two M responses.

V.—Discussion of Results.

There seems to be no doubt that the constructive use of sand can be considered a sign of good intelligence. The experiment also supports the view that there is a relationship between this sign and that of the Rorschach M response, implying that the diagnostic import of sand construction is not only that of good intelligence, but indicates also a qualitative personality difference, namely the availability of inner resources of imagination.

Imaginative intellect may serve most often as a stabilising influence, but can, of course, go beyond what is optimum for social adjustment. One of the finest Worlds for sand construction was made recently in a hospital by a paranoid schizophrenic. She made a vivid mountain scene in the Lake District, where a car was shown crashed against a tree. She said that this scene was herself, " very bleak ," and that it was also her childhood home. Then she made an expressive (peaceful) valley scene with a cottage, which she described as her ideal self, in contrast to the actual bleakness. In the same way a Rorschach protocol from certain kinds of schizophrenics may show a very high M column and no balancing FC at all, where the individual has abandoned ordinary life in favour of fantasy.

On the other hand the student who ignored the sand, and who tends to have skin trouble when she is emotionally upset, is clearly of a somewhat hysterical personality. She gave M responses in the Rorschach test, but perhaps in face of the World task, with its more life-like materials, ' inexorable repressions ' prevented her from using the sand constructively.

There is a suggestion from these individual results and others, that the constructive use of sand is more associated with introversion than with extraversion. This is to be expected from the implications of the M column in the Rorschach.

The control group showed interesting contrasts, between individuals who expressed themselves factually, e.g., they depicted an actual place and pointed out its details, and other individuals who expressed themselves symbolically, e.g., they made a scene and explained that it had further meaning, as for instance ' the spirit of Pakistan.' Dr. Buhler* once formulated the hypothesis " that the threshold for the disposition to produce and to perceive symbols is individually a very different one." Many schizophrenics for example seem to have extremely low thresholds, and it may be that the secret of getting into rapport with such persons is to be not too dissimilar in this respect. The sand formation was frequently the aspect of the World chosen by members of the control group for symbolic expression. This would provide a standard situation for attempting research on the extent of individual and group differences in the disposition for symbols among normal children and adults. Such differences may be important for inter-personal understanding in the social and educational fields.

---

* Buhler, Charlotte. Symbolic action in children. *Transactions of the New York Academy of Sciences*, No. 17, 1941.

SUMMARIES OF RESEARCHES REPORTED IN DEGREE THESES

Co-Education and Attitude Towards the Use of Force.

By G. P. BEAUMONT

(Summary of thesis presented for the degree of M.A.(Education) of the University of London. 1957.)

The object of this research was to investigate the effect of co-education on attitude towards the use of force. This attitude had already been studied by R. Stagner (1944), who showed that it correlated positively with other aggressive attitudes. Relevant also was the work of H. Eysenck (1954), which discusses an analogue of James' conception of Tough and Tender-mindedness. From a constellation of attitudes, Eysenck isolated a primary factor, T, which makes its appearance "essentially as a projection on the field of social attitudes of certain fundamental personality traits." Tough-mindedness was shown to be associated with aggression and dominance.

The construction of a Thurstone equal-appearing-interval scale, on the basis of 105 items, nine categories, and twenty judges was attempted. The content of the items frequently referred to those aggressive uses of force which have an ambiguous position, in a competitive society.

Examples :

Item 100—Only a coward funks a fight.
Item 39—What you don't take, somebody else will.
Item 86—Every soldier who kills is a murderer.

Hovland and Sheriff (1952) having re-opened the question of personal bias in the item judging process, the work of the rejected judges was scrutinized and found not to constitute a homogeneous sub-group.

After item elimination by inspection and the calculation of median scale values, gaps in the middle of the scale led to the adoption of a five-category Thrustone-weighted system. A 'higher' score will imply relatively greater disapproval of force. Fifty items were administered to a pre-test group of fifty-two boys and fifty-two girls. After an empirical check on item discrimination, a final set of twenty-seven items was found to have a split-half reliability of ·89. Girls scored significantly higher than did boys, on both halves.

The final twenty-seven items were administered to all available fifth-form pupils in a girls' high school, a boys' grammar school, and a co-educational grammar school, all under the same L.E.A. The results are contained in Table 1.

TABLE 1

	Non-co-ed.	Co-ed. boys	Co-ed. girls	Non-co-ed.
Sample size	49	30	36	55
Sample mean	145·5	143·5	149·8	155·6
Corr. s.s.	8746	3446	5515	10596

The 'F' test revealed no significant ratios of estimated variance. Application of the 't' test gave the results contained in Table 2.

TABLE 2

Difference	d.frdm.	't'	significance
Non-co-ed. *vs.* coed. girls	89	2·01	5% level
Co-ed. girls *vs.* coed. boys	64	2·15	5% level
Co-ed. *vs.* non-co-ed. boys	82	—	not significant

It was remarked that the girls in the co-educational scored significantly lower than girls in the non-co-educational school.

Parametric methods of analysis make fairly severe assumptions about the distribution of the scores and the nature of the metric from which they are derived. Non-parametric tests, notwithstanding their relative inefficiency, have the advantage of assuming only similarity of distribution in random sampling and are frequently easy to apply. Moses (1952), has described the 'Median' test, which consists of pooling the scores from two samples, which can have different sizes, and counting the number of scores from each sample above and below the joint median. Scores falling at the median can be shared out or eliminated. The frequencies may then be entered in a two-by-two contingency table and the goodness-of-fit tested by a Chi-squared test.

The application of the 'median' test failed to reveal any significant difference attributable to Co-education but the pooled girls' scores remained very significantly higher than the pooled boys' scores.

Non-parametric methods are frequently applicable where the metric may be deemed to be only an Ordinal scale. In this latter connection, a fairly comprehensive review of Scalogram analysis was carried out. The central concept of Scalogram analysis is: "if a single quantitative score is to represent, without ambiguity, the behaviour of an individual on a group of items, then it must be possible, knowing the individual's score, to know his behaviour on each and every item in the group." The coefficient of reproducibility, defined by Guttman (1940), which measures the extent to which this is possible, is very sensitive to the modal response.

From the schedule of twenty-seven items, eight items were administered to a new sample of eighty-three pupils. The items were those which discriminated well between high and low scorers and were scored with two categories, agree and disagree.

Modal reproducibility coeffcients for girls ahd boys were 69 per cent. and 66 per cent. respectively. Using the 'Cornell technique,' the corresponding maximum reproducibility coefficeients were 84 per cent. and 79 per cent. Non-scalability was inferred.

The tentative suggestion arising from this limited study is that the girls in the co-educational school gave significantly more approval to the use of force than did the girls in the non-co-educational school.

REFERENCES

Eysenck, H. J. (1954). *The Psychology of Politics.* Routledge and Kegan Paul.

Hovland, C. I., and Sherif, M. (1952). Item displacement in Thurstone scales. *J. Abnorm. Soc. Psychol.*, 47, 822-832.

Moses, L. E. (1952). Non-parametric tests for psychological research. *Psychol. Bull.* 49, 122-143.

Stagner, R. (1944). Studies of aggressive social attitudes. *J. Soc. Psychol.*, 20, 109-140.

Stouffer, Guttman, *et al.* (1940). *Measurement and Prediction.* Vol. 4, Princeton.

An Examination of the Causes of Deterioration in Academic Performance Among Pupils in a Grammar School.

BY S. GRIFFITHS

(Abstract of thesis submitted and approved for the degree of Master in Science in the University of Wales, Swansea.)

AIMS OF THE INVESTIGATION.

The investigation seeks the causes of serious deterioration in the academic performance of individual pupils. The survey was made in the catchment area of a co-educational secondary grammar school of over 500 pupils. It is a case study enquiry attempting to supplement such large scale statistical investigations as *Early Leaving* (Ministry of Education, 1954), and *Social Class and Educational Opportunity* (Floud, *et al*, 1957).

PLAN OF THE INVESTIGATION.

(*a*) Survey of home background conditions of academic deteriorators in their first year at the grammar school, between 1950 and 1954 (inclusive). This section involved thirty-nine case studies, twenty-one boys and eighteen girls.

(*b*) A socio-economic survey of the environment of one junior school in the grammar school area ; and an investigation of the academic progress of all pupils from that junior school who entered the grammar school during the five years covered by the thesis. This particular junior school was chosen because it produced a higher proportion (relative to the other junior schools in this area) of academic deteriorators.

(*c*) Subsidiary exercise : Investigation into the home background of pupils who showed conspicuous academic improvement at the grammar school.

PROCEDURE.

Deteriorators were found in the first instance by the criterion of demotion ; but two statistical methods were devised to check its validity. Use was made of these statistical methods to discover ' improvers ' from one year's entry ; case studies of these ' improvers ' were compiled.

Causes of deterioration were sought by case study. These studies included reports on the attitudes and performance of each deteriorator for the last year at the junior school, results of the relevant annual entrance examinations, the opinions of school staffs, social workers, the chief medical officer for the county and the parents of the pupils concerned. The deteriorators were given attainment and intelligence tests (the latter were given to the whole year's entry) ; and their performance in these tests was compared with that of non-deteriorators and ' improvers.' The attitude of the deteriorators to subjects in the curriculum (at the grammar school) was compared to that of non-deteriorators to the same subjects. Reading and Arithmetic Quotients were obtained.

Home background was investigated under the following headings : facilities for doing homework, academic culture in the home, financial conditions and number of children. The educational background of the parents (that is, if they had been educated at a grammar-type of school) and the occupation-category of the parent(s) were ascertained ; also the attitude of the parents to education in general and the education of their own children in particular was sought. The attitude of the parents of the ' improvers ' to education and the academic progress of their children was similarly investigated.

The ' personal ' attribues investigated were : intelligence, temperament, emotional disturbance, interests (especially out of school), physical defects (for example, slight deafness, minor eye-sight trouble).

Investigation was made into the following aspects of the problem : conditions of work in the junior schools especially as regards preparation for the entrance examination to the grammar school ; extra-mural coaching for the entrance examination ; organisation of the curriculum and supervision of the entrants during their first year at the grammar school ; absence from school ; long distance travel to and from the grammar school; choice of friends in and out of school hours; emotional upset at changing schools; loneliness at the grammar school.

An attempt was made to estimtae the relative influence of the factors investigated on academic deterioration. The case study of each deteriorator was analysed and the factor which appeared to have had the greatest effect for deterioration noted in each case. (See Table 1.)

RESULTS OF THE INVESTIGATION.

(1) *Factors of Influence.*

TABLE 1

COMPARISON OF SOME FACTORS CAUSING ACADEMIC DETERIORATION.

Factor	Boys (N=21)	Girls (N=18)
Unfortunate home background	14	11
Personal Qualities	6	4
Other causes	1	3

The numbers dealt with in this Table are very small; but *Early Leaving* also stresses the major influence of home background on the academic progress of grammar school pupils.

Under the heading 'Personal Qualities,' lack of persistence at school work was the predominant factor.

(2) *The Eduucational Background of the Parents.*

In only one case amongst the thirty-nine deteriorators in the investigation was there a parent educated in a grammar school, whereas an enquiry into the parental education of all pupils in two years' entries showed (Table 2) that, in 108 families out of 181, one or both parents had been taught in a grammar-type of school. (Brothers and sisters were allowed for in Table 2.)

TABLE 2

THE EDUCATIONAL BACKGROUND OF PARENTS.

	'A' stream		'B' stream		'C' stream	
	B	G	B	G	B	G
No. of children	30	34	27	36	24	30
(*a*) Only mother attended gr. sch.	7	3	5	7	6	3
(*b*) Only father attended gr. sch.......	12	10	8	9	1	3
(*c*) Both parents attended gr. sch.	8	13	3	8	1	1

On the other hand, at least one parent of every 'improver' investigated had grammar-school education. Again, we must emphasise the snallness of our numbers; but the same pattern is found in *Social Class and Educational Opportunity*, where it is pointed out that these better educated parents are more encouraging to their children's academic progress than are the less educated parents.

(3) *Parental Occupation Categories.*

In addition to ascertaining the occupation categories of the parents of the deteriorators, the occupational groups of the parents of all the entrants in one year were obtained. Table 3 compares the occupational categories of the deteriorators, 'improvers' and others.

TABLE 3

CLASSIFICATION FOR SOCIAL STATUS, FROM PARENTAL OCCUPATION.

Category	1	2	3	4	5	6	7
No. of Pupils (one year entry) . .	6	9	10	9	41	21	6
No. of improvers (over 3 years) . .	—	4	5	5	7	—	—
No. of deteriorators (over 5 yrs.)	—	—	—	2	15	12	10

The classification for social status from parental occupation is based on *Social Mobility* (Ed. D.V. Glass, 1954) ; the categories are as follows :

(1) Professional and High Administrative.
(2) Managerial and Executive.
(3) Inspectional, supervisory and other non-manual.
(4) As for 3 but lower grade.
(5) Skilled manual and routine grades of non-manual.
(6) Semi-skilled manual.
(7) Unskilled manual.

It will be seen from Table 3 that the deteriorators tend to congregate in the unskilled, semi-skilled and skilled groups. It will also be noticed that there is a very high percentage of deteriorators in the unskilled category in view of the relative smallness of entry in that category. Both *Early Leaving* and *Social Class and Educational Opportunity* draw attention to the academic failure at the grammar school, of a large percentage of the children of unskilled and semi-skilled parents.

SUMMARY OF CONCLUSIONS.

(The figures in brackets indicates the number of deteriorators out of thirty-nine affected by the factor.)

1.—A home which cannot provide facilities for quiet study in comfort is an unfortunate one for grammar school pupils (22). On the other hand, a home which provides facilities is an asset as found in the case of the improvers.

2.—The parents of the deteriorators lacked grammar school-type of education (38).

3.—The majority of the parents of the deteriorators were in the lower categories of occupation (22), whislt those of the 'improvers' were in the higher categories.

4.—The deteriorators had less encouragement from their parents to do well at school (37) than did the improvers from their better educated parents.

5.—Whilst the deteriorators (38) had a persistence factor for school work below average (estimated by their teachers), that of the improvers was above average.

6.—Frequent absences from school for inadequate reasons is a contributory cause of deterioration (10).

The Influence of Home Television Viewing on Children in a Mixed Secondary Modern School.

By W. W. MITCHELL.

(Absrtact of thesis submitted for the degree of M.Ed. in the University of Manchester, September, 1958.)

This investigation was carried out between January, 1957, and July, 1958, in a mixed secondary modern school in Cheshire, having 300 to 400 pupils from urban and rural areas.

The following questions were investigated :

(1) The amount of time the children spent viewing on weekdays, with differences according to age, sex and ability ; and an evaluation of this time based on an arbitrary assessment of the programmes seen.

(2) The effect of viewing upon homework and the distribution of out-of-school time on weekdays.

(3) The effect of viewing on reading, hobbies, pastimes and interest in the theatre and orchestral concerts.

(4) Television viewing as a source of general knowledge.

(5) The effects of viewing adult plays from the ' legitimate theatre,' such as Arthur Miller's " Death of a Salesman."

(6) The opinions about the effects of viewing held by the children themselves and by their parents.

The methods used for the above sections were respectively :

(1) A survey was made of all the viewing done by the school on three weekdays in January, 1959. During two school days following, pupils indicated from a numbered list the proportion of each B.B.C. or I.T.A. programme they had viewed, and from these replies the various categories of viewing amounts were calculated. This was repeated with a small sample for two days in the Summer. Other information concerning general viewing frequency, preferences for B.B.C. or I.T.A. programmes, etc., was obtained from special items on questionnaires used in (3). Each programme was categorised by a panel of teachers according to its value as entertainment or information and from the numbers viewing each programme was calculated the total school viewing time for each type of programme.

(2) Forms doing homework recorded exact times spent on homework and on viewing on two test evenings and also answered questionnaires concerning the influence of viewing and other activities on their homework.

(3) Questionnaires were answered by all pupils in 1957 with a follow-up questionnaire on selected items one year later in 1958.

(4) An objective test representing a reasonable amount of knowledge about animals was constructed and replies were correlated with viewing frequencies.

(5) Questionnaires concerning opinions, interest and understanding were completed by small groups of children who had viewed selected plays.

(6) In the classroom, children wrote free expressions of opinion about television viewing, and a proportion of parents did the same voluntarily in reply to a circular letter.

The main conclusions of the work were :

(1) Nine-tenths of the children tested viewed television regularly for an average of 2 hours 45 minutes each weekday. The A forms viewed significantly less than the C forms but there were no differences by sex or age.

(2) I.T.A. programmes for children and for adults were preferred to B.B.C. programmes although this was less marked with girls than with boys.

(3) Seven out of ten children viewed some part of the children's programmes on each weekday tested.

(4) Half of the children never listen to B.B.C. sound Children's Hour and those who do so view B.B.C. television programmes more than I.T.A. programmes.

(5) Two-thirds of the viewing time of the school was devoted to programmes whose main purpose was to entertain rather than to instruct.

(6) Most children gave the required time to homework tasks and these were mostly completed before any viewing was done. Although viewing was reported as causing interference sometimes with the proper completion of homework, it did not do so any more often than other named causes.

(7) There was a slight negative relationship between viewing and amount of reading out of school.

(8) Viewing caused some stimulation of interest in hobbies, etc., with a minority of children.

(9) There was no evidence that viewing had stimulated any positive interest in the theatre or in orchestral concerts among the children.

(10) Although television is the most common source of information for children about animals, the results of an objective test of this subject showed no relationship with the amount of viewing.

(11) Children themselves feel that television viewing brings many advantages principally in the form of information and widening of interests, but at the expense of much free time and initiative.

(12) Children are often capable of making appreciative and discerning responses to television performances of plays written for an adult audience in a legitimate theatre.

(13) The parents who replied felt that viewing has created many new difficulties for their children and their homes but that its ultimate effect depends on the attitude and the actions of the parent.

An English Attainment Test for Blind Children.

By JEAN M. MURRAY.

(*Abstract of thesis submitted for the degree of M.Ed. in the University of Manchester, September*, 1958.)

The aim of this investigation was the construction of an English attainment test for blind children that would satisfy the following criteria: it should be objective and reliable, it should not require any knowledge of braille and it should be suitable for as wide an age range as possible.

To be independent of braille it would have to be either an individual test or a group test administered orally, with some device other than braille for recording responses. The construction of an individual test appeared to be a relatively straightforward task but, as the test would probably be used by the children's own teachers, this had the disadvantage that the results would be liable to the influence of teacher-child relationships and halo effects. It was, therefore, decided to explore the possibility of a group test, and after some experiment it was found that multiple-choice answers could be recorded on the Taylor frames that the children normally use for arithmetic.

A try-out test of 150 items was constructed and given to all the children between the ages of 9 and 16 years in the blind schools in England, results being obtained from 516 of the 592 children in the age range. From this try-out test, fifty-five items were selected by item analysis to form the final test which is suitable for children between the ages of ten and sixteen years, though there is some doubt about its discrimination for the ten-year-old group. The test is preceded by a practice test to familiarise the children with the method of recording responses, and the whole test takes between 50 and 60 minutes to administer.

This test has a mean facility of 56·8, a mean discrimination (upper and lower thirds) of 0·564, and reliability coefficients of 0·949 (Kuder-Richardson) and 0·944 (split-half). There are no significant sex differences. The English scores show a correlation of 0·611 with intelligence quotients (Williams Intelligence Scale for Children with Defective Vision) for a representative sample and of 0·885 with the mental ages calculated from these I.Qs. The correlation between English score and chronological age is 0·573. The mean scores for successive one-year age groups show a steady rise with age. Tables of age norms, as an aid to teachers in interpreting scores, have been obtained graphically from these means.

The validity of the test has been difficult to assess as there are no other English tests for blind children with which to compare it. Using comparable tests for sighted children as the criterion, the content validity of the test is satisfactory though there are certain omissions (e.g., reading comprehension) that are inherent in the concept of a test for blind children that assumes no knowledge of braille. When the scores were compared with teachers' estimates and the rank order correlations were combined, a mean coefficient of 0·62 was obtained. The correlations between English score and chronological age, mental age and intelligence quotient were of the expected orders of magnitude. Thus, the qualitative evidence of content validity is supported by such quantitative evidence of concurrent and construct validity as is available. The predictive validity of the test has not been investigated.

No attempt has been made to discriminate between children with different degrees of blindness, or between congenital and adventitious blindness, or to make allowance for different lengths of schooling.

CRITICAL NOTICE

BY CYRIL BURT.

H. T. HIMMELWEIT, A. N. OPPENHEIM and PAMELA VINCE (1958). *Television and the Child*. London : Oxford University Press, pp. xxii + 522, 42s.

Every new medium of communication has, in its time, set the public arguing whether its harmful effects may not outweigh the good. The cinema, the radio, and even reading, have each been blamed for the laziness, the backwardness, and the petty delinquencies of the young. Now it is the turn of television. Hitherto, however, there has been little or no factual evidence to supoprt the allegations so commonly made. How much time is actually spent gazing at the television screen, and what types of programme are preferred ? What kind of youngster turns into a television addict, and what are its effects on school work, family life, physical health, emotional stability, and character ? These are the questions which Dr. Himmelweit and her colleagures have set out to investigate. The research was sponsored by the Nuffield Foundation ; and throughout their inquiries they received the help of the Audience Research Department of the B.B.C., and an advisory committee of statistical and psychological experts.

Their investigation consisted of two principal surveys. In both they concentrate primarily on two age groups—pupils between 10 and 11 from the top classes in the primary schools, and pupils between 13 and 14 from grammar and secondary modern schools. The ' main survey ' was carried out in the Summer of 1954 in four English cities, London, Portsmouth, Sunderland, and Bristol ; here, from a total of 4.500 children all systematically tested, they selected 927 ' viewers ' matched for sex, age, intelligence, and social background with an equal number of ' controls.' At Norwich the opening of a new television transmitter supplied an opportunity for a second survey consisting of a ' before-and-after study ' : out of the 2,200 children tested in 1955, before television was available, they were able a year later to trace as many as 185 who had television sets, together with an equal number of matched controls. Every child was asked to keep a written diary for one week, stating what he did each day from the time he left school up to bed-time. Six weeks later the children were given a questionnaire on their various interests and leisure activities, with special reference to reading, the radio, the cinema, and television. Teachers furnished ratings for their pipils' abilities, attainments, and general behaviour.

The data obtained are summarized in seventy-three tables in the appendices ; and the main conclusions are discussed under six main headings. The report begins by describing the nature of ' the child audience.' As compared with the control group, the ' viewers ' are " more extraverted, more eager for ready-made entertainment, and less mature in tastes and mental development " ; and in the main come from homes where supervision is more relaxed. On an average they spend eleven to thirteen hours a week watching television—far more than is spent on any other leisure activity, but far less than would be spent by children of the same age in the United States. Among younger children television has greatly reduced attendances at the cinema ; but many of the older children, particularly the girls, still prefer the cinema : " you can sit in the dark with boy friends, you can't have at home," and " you can see your favourite film-star." In two-thirds of the homes the set is left going throughout the evening.

What kind of subjects do the viewers prefer ? Less than a third of the time is devoted to programmes designed for children. Subjects such as hobbies, animal programmes, and documentaries generally, attract little attention. With girls as well as boys ' crime thrillers ' take first place. ' Fabian of Scotland Yard ' was the favourite. The other items selected—' The War in the Air,' ' Ask Pickles,' Westerns, and comedy scenes—differed with age, intelligence, and individual interests. These preferences indicate the reasons for television's strong appeal : it offers constant change, excitement, and suspense, and opens up an easy escape from everyday chores into a fictitious world of glamour, light-heartedness, and romance, where each can identify himself with his favourite hero. No programme was mentioned by more than a third of the children. Thus, even among those of the same age and intelligence, variation in taste is unexpectedly wide. Where only one channel is available, a few children hit on highbrow topics—' Science Review,' ' Meet the Commonwealth,' ' Have you a Camera ? '—and enjoy them. But, when the child can follow his preferences by switching from one channel to another, he is less likely to turn to an educative or informative programme.

Addiction to television is a symptom rather than a cause. The addicts are chiefly drawn from the less intelligent boys and girls. They consist, we are told, very largely of those who are in some degree maladjusted—shy, retiring youngsters who would seek refuge in any kind of cheap entertainment that happens to be available. Without a television set, they would frequent the cinema : without the cinema, they would pore over comics or the cruder types of childish thriller.

What effects then does television have on the child's sense of values and social outlook ? The world presented, so the investigators believe, is predominantly that of an upper middle class urban community. The occupations of such people are usually depicted as worth-while, and manual work as dull and forbidding. As a result, the viewers prove to be far more ambitious in their aims. For the rest, life is portrayed as difficult and problematic, especially for the feminine sex ; marriages are usually unhappy ; relations between parent and child are strained. The moral is that virtue alone is not enough : toughness and self-confidence are needed for success.

The liability of television to provoke neurotic disturbances or delinquency, has, it would seem, been greatly exaggerated. Fisticuffs and fighting arouse fear only when they occur in programmes purporting to describe actual life, not in fictional dramas. Knives are more alarming than guns ; verbal aggression, ridicule, and reprimand more disturbing than either. Those most affected are the very young and the emotionally insecure. Their anxieties and subsequent nightmares are enhanced by viewing in complete darkness, particularly when no adult is present.

The gain to general knowledge appears comparatively slight. Even when the child picks up useful facts, he could have done so far more rapidly and accurately had his time been spent on books. "A net profit was demonstrable only in the case of the younger and the duller children ; and even these carry very little over into their performances at related subjects taught in school." Nevertheless, in scholastic attainments the average viewer seems to hold his own as compared with non-viewers of similar age and intelligence ; but in both age groups the more intelligent are apt to fall a little behind.

The possession of a television set tends to keep the family together, with the children under the parental eye. Conflicts may arise over meal-time, bed-time, and the interruption of viewing to run errands or perform household tasks. Sometimes television is used as a means of reward or punishment. But few parents recognize any need to control the content or the actual amount of viewing.

The effects of television on the childs' leisure activities seem for the most part to be indirect rather than direct. The most important consequence is that the child simply has less time for the things he might otherwise do. He spends fewer hours in the open air or on out-door games ; and goes to bed about half-an-hour later. There are, however, no indications that viewing impairs physical health or induces headache or eyestrain, as is so often alleged.

On the main issues, it will be seen, the upshot of the whole inquiry seems at once conservative and reassuring. " Television is not as black as it is painted, neither is it the harbinger of culture and enlightenment which its enthusiasts have sometimes claimed : its capacity for influencing the child's horizon is not spectacularly different from that of any other mass medium..'" The book concludes with a series of recommendations addressed to parents, teachers, and producers, and concrete suggestions for further research.

Parents are urged to give better guidance in regard to the choice of programmes and to set a better example in their own selection. This, it is added, is especially important in regard to programmes intended for adults which the older boys and girls nearly always watch. They should, moreover, discuss the subjects with their children both before and after viewing. Such counsel is admirable in itself. Unfortunately, it is members of the less intelligent families who most need advice, and are least likely to get it. Except, therefore, in the more enlightened homes, it would seem that the necessary guidance will have to come from outside.

The suggestions to the teachers have a far better chance of being carried into effect. Most of the teachers had definite views on the matter. Only 12 per cent. thought home-viewing undesirable, chiefly, we are told, because it tends to interfere with homework. But, if I may judge from my own experience, the objections go a little further than this. Many teachers complain that, not only the songs, the comedies, and the Westerns, but also the more serious programmes intended specifically for children are uttered in the highly infectious accent, idiom, and vocabulary of the more vulgar transatlantic type of film. And secondly, it is said, excessive viewing tends inevitably to develop an inert and passive attitude of mind, and to encourage a torpid desire to have everything dished up in a picturesque and diverting fasion, with the result that any form of verbal presentation, whether in the classroom or through books, becomes an insufferable bore. However, in the answers of the children themselves, Dr. Himmelweit and her colleagures could discover no factual evidence to justify such complaints—a point in which their findings apparently differ from those of Dr. Belson in his studies of adults. Among the teachers, too, as many as 15 per cent.—mainly from primary or secondary schools—reported that television had created fresh interests—particularly in such subjects as history, geography, and nature study. Another 9 per cent., mostly from infant or junior schools, thought it had enriched the children's vocabulary and given them something to talk about. Many, however, realized that much more could be done to enlarge the benefits and diminish the risks.

The main recommendations put forward by the investigators is that teachers should examine the published announcements of forthcoming programmes, and point out to their pupils what they should try to watch. But without far more detailed advice and guidance the teacher's judgment is not likely to have much value. A beginning might be made in training college courses. Practically all give instruction on the use of television in the school. But, with one notable exception, hardly any take up the more elusive problem of television in the home. The Society of Film Teachers might usefully extend its scope to include television. The joint committee set up by the Association of Teachers in Colleges and Education Departments (in conjunction with the British Film Institute) has already given much attention to school programmes on television, and might easily include the wider topics of viewing in the home.

Finally, the investigators make a number of useful suggestions which deserve the serious study of producers. As they point out, the distinction drawn in the printed programmes between 'Children's Television' and the rest is in practice quite ineffective. Instead, they suggest, the *Radio Times* and the *T.V. Times* should specifically indicate what programmes are suitable for children below such and such an age. I would add one further point: with classical productions, such as children should ordinarily be encouraged to watch—*Jane Eyre*, *Wuthering Heights*, *Our Mutual Friend*, and the like—both producers and actors tend to magnify whatever is sensational: in order to secure theatrical effects they make the most of the crimes, the hysterical outbursts, the fights, the panics, and the love-scenes—without quite realizing the reactions they are likely to evoke in the children. But on all these points there could with advantage be far more consultation and co-operation between producers, teachers, and educationists.

As the investigators rightly insist, each of the many problems which their surveys have raised calls urgently for further research. Their own investigation is not only the first of its kind, but in the main admirably planned and assiduously carried out. Its great merit is that, both in the general scheme of the inquiry and in the discussion of the conclusions to be drawn, the emphasis is squarely laid, not simply on the possible effects of one particular medium of communication, but on the attitudes, the disposition, and the social outlook of the children themselves.

BOOK REVIEWS

NATIONAL FOUNDATION FOR EDUCATIONAL RESEARCH IN ENGLAND AND WALES. *Educational Research*, Vol. 1, No. 1, Nov., 1958. London: Newnes, issued thrice-yearly, 5s. 6d. each issue.

Most readers are probably already aware of this new venture; however, this *Journal* is glad to offer some words of avuncular welcome to a lusty infant, albeit the infant threatens to grow up into a business rival! The professed aim is to interpret educational -psychological research to the teacher, in relatively non-technical terms, rather than to describe new researches as such; and the first number includes several admirable articles of this type such as J. M. Morris on phonics in reading, Yates and Pidgeon on selection at 11+, and Wall on motivation. But one wonders how the supply of such surveys is to be maintained to the tune of 140,000 words a year. Already, the longest article, by Pidgeon, describes in detail a comparison between the arithmetical attainment of Californian, Australian, and British 11-year-pupils—a topic which is obviously of more interest to, say, our own readers than to teachers in general. However, there is plenty of room for two publications, reaching somewhat different audiences, even if it turns out that they overlap rather more than the Editors believe.

The National Foundation's journal is easy on the eye; both tables and bibliographies are clearly printed. The page size is such that it will stand neatly on the same shelf as the *British Journal of Educational Psychology*. May we hope that our yellow covers and the new red ones will accumulate in close proximity in all training college and university, and many private, libraries. P.E.V.

AINSWORTH, M. D., and AINSWORTH, L. H. (1958). *Measuring Security in Personal Adjustment*. London: Oxford University Press; Toronto: University Press, pp. xiii+48, 24s.

This report describes a part of an extended programme at the Institute of Child Study, Toronto, in which an attempt is being made to devise a series of tests that will measure the extent of a person's security or insecurity in five major areas of adjustment: intimate family relations, interpersonal relations outside the primary family, vocation and the preparation for it, leisure activities and philosophy of life. The tests so far developed deal largely with four of these areas (omitting vocation). Each test consists of four, or where appropriate five, sub-tests which are designed not only to measure security, but also the ways by which a person attempts to relieve tension when feelings of insecurity result from inability to satisfy some pressing need. The kind of solution adopted by the individual, and the measure of its use, reflect the extent to which insecurities in early years of development have prevented growth towards secure interdependent relationships which typify adult adjustment. On this basis the sub-tests, each consisting of six to twelve statements, cover Mature Dependent Security, Independent Security, Immature Dependent Security, Insecurity, and Deputy Agents (Defence Mechanisms).

The present book outlines the theory of personality development from which the rationale of the tests is derived, and gives an account of the development of the sub-tests through several revisions, illustrating the "constant process of checking back and forth between tests and rationale and between rationale and theory." The results of item analysis, and the inter-correlations between sub-tests, which showed the extent to which the obtained relationships were consistent with expectancies in terms of the underlying rationale, were used throughout this process. Detailed results are reported for the final revisions. The conclusions draw attention to aspects of the tests which might be improved and discuss the advisability of using the tests, as a whole series, or in part, on the basis of their reliability and validity as so far demonstrated. While the tests were devised for young adults of college age, with some exceptions they appear to be suitable for wider application. The authors recognize the need to repeat these studies with different subjects in order to check the reliability of the intercorrelations upon which the present case for validity rests.

This is a valuable piece of work, reported with considerable care and unusual clarity. It has particular value in that it presents, in a most difficult area of test construction, a model in which personality theory and statistical analysis have been brought into a working relationship. The result is a test that has a theoretical basis and a rationale which are widely acceptable, and which lend themselves to extension and refinement as the test is further developed. H. PHILLIPSON.

BERNHERT, ELEANOR H. (1958). *America's Children.* New York: John Wiley and Sons; London: Chapman and Hall, pp. 185, 48s.

This is a statisical analysis in the Census Monograph Series based on the 1950 Census of Population in the United States. It deals mainly with the education of children in relation to their social and economic backgrounds. Chapers deal with the number and distribution of children in the U.S.A., the size of the family, the numbers of broken homes, comparisons of Southern, rural-farm and non-white school standards and employment and its effect on early leavers. Many of the facts reported have their counterparts in this country but direct comparisons are difficult and perhaps misleading. Retardation in school performances and scores on a general literacy test had a high positive correlation ($r = \cdot 92$) with poor housing conditions, and with overcrowding ($r = \cdot 80$). In 1950, about 51 million (roughly 34 per cent.) of the population were 19 or under 19 years old and of these over four million were members of broken families; here widowhood and marital discord were the two main contributory factors. Attainment in school was measured by the grade (or form) reached by the children and variations were found in retardation and acceleration between white and non-white, rural and urban, boys and girls. The problem of retardation is dealt with in some detail. It is estimated that 25 per cent. of the high school population was at least one grade below the average and that about 50 per cent of the high school students (i.e., 80 per cent. of the total agegroup) dropped out before completing the course.

With the large increase of child population after the last war, the U.S.A. is faced with the problem of the 'bulge.' The numbers enrolled in high schools is expected to increase by about 3 per cent. annually up to at least 1964, with the peak occurring in 1960. W. H. KING.

BLACKWELL, A. M. (1958). *Lists of Researches in Education and Educational Psychology, Supplement III.* London: Newnes, pp. 64, 15s.

This supplement, covering theses of 1956 and 1957, follows the same pattern as its predecessors, and is only slightly longer than the 1954-5 volume. We must recognise yet again our great debt of gratitude to Mrs. Blackwell and the National Foundation for Educational Research, though qualifying our thanks once more with a request for an amalgamation of the five volumes.

CLARKE, A. M., and CLARKE, A. D. B. (edit.) (1958). *Mental Deficiency: The Changing Outlook.* London: Methuen, pp. xvi+513, 45s.

For far too long mental deficiency has been regarded as a medical problem, and as medical treatment was generally ineffective, defectives have been left to vegetate—like the incurably diseased—in hospitals and institutions. Among the 800 odd references in this book, few names of psychologists are to be found, except Binet's and Burt's, until the 1950s. However, the work of Kirk and Sarason in the U.S.A. heralded a change, and since the war a remarkable group of British psychologists—the Clarkes, Tizard, O'Connor and Gunzburg—have revolutionised our outlook, and they are the chief contributors to this important compendium on the psychology of the defective and his psychological rehabilitation.

There are excellent, up-to-date chapters on incidence, classification and etiology, but the main weight is on the measurement of intelligence and its modifiability, on diagnosis (including the clinical application of projective techniques), on learning, educational and vocational training, and social adjustment. The tremendous variety of defectives is also well brought out: there is no more excuse for thinking of them as a single group occupying the tail end of the distribution curve than there is for thinking of physical ill-health as a unitary character. In addition to the authors mentioned above, Elspeth Stephen provides a detailed account of Cerebral Palsy and Mental Defect, and Margaret Penwill a chapter on Speech Disorders and Therapy in Mental Deficiency. For the most part the writers adopt a common standpoint, but there is a certain amount of unnecessary overlapping, and an occasional lack of balance. For example, O'Connor hardly convinces us that the most significant contributions to our knowledge of learning among defectives emanate from Russia ; and A. D. B. Clarke goes into excessive detail in describing the admittedly important but already published, work of the Social Psychiatry Unit on motivation of imbeciles. But it would be ungrateful to cavil because the book is somewhat less compact and readable than it might have been. Throughout, it is extremely clear, rationally argued (apart from a few incautious statements) and well-documented ; and throughout attention is drawn to the many gaps in our knowledge.

To this reviewer, at least, the book is welcome not only because it represents a new advance in psychology, but because it provides a sound basis for more enlightened handling of feeble-minded and imbecile children and adults, and thus for the reduction of wastage and human misery. P.E.V.

FLEMING, C. M. (1958). *Teaching : A Psychological Analysis*. London : Methuen, pp. 291, 28s.

The amount of psychological information relevant to the teacher's work is now so extensive and complex that it is very doubtful if it may be compressed into one book. But, if the attempt is made—and Dr. Fleming's recent book exemplifies it—one may judge it by three basic criteria : critical selection from the vast array of material available, compactness of presentation, and organisation.

As one with a knowledge of Dr. Fleming's previous work would expect, this present book is in effect an encyclopaedia of information on a multitude of psychological topics bearing on educational practice. For such a compilation she merits the admiration and gratitude of those who have to teach psychology to teachers, but who may not have the experience, skill, time or energy to undertake such a task themselves. The criterion of compactness (*multum in parvo*) is well met.

But, as one with a knowledge of the difficulties involved, would expect also, the exploitation of this one criterion has resulted in a less satisfactory regard for the other two. Of the topics one would like to see included, few have been omitted ; but the book is so compact that each topic receives only a flying visit. The reader is dragged round the psychological sights in record time, and while nearly all the exhibits are worth seeing, yet he has no time to see them properly. Such treatment may perhaps satisfy the psychologist who has seen them all before ; but it is doubtful if it will appeal to the teacher or, indeed, if it is fair to the exhibits themselves. Some topics of course receive rather more space than others ; and in the sections on Child Development, for example, one can enjoy for a page or two the luxury of a sustained argument and the elaboration of a single explicit theme. On the other hand, some issues are given such scant attention or dismissed so summarily that it would be quibbling to say they had been selected at all. (The views of Guthrie and of Pavlov, for instance, are not nearly so simple and invalid as Dr. Fleming's cursory rejection of them suggests.) At the present stage of psychology's development, the succinct generalisation demanded by this kind of book is by no means easy, sometimes impossible, and likely to be misleading.

As for the organisation of the subject-matter, those who have given serious thought to the problems of teaching psychology to teachers surely will agree that Dr. Fleming's plan was the right one. This plan, consisting essentially in a structuring of the relevant material in accord with the educational (rather than purely psychological) context in which it is to have pedagogical significance and be applied, has of course, been put into operation before—notably by Cronbach in his *Educational Psychology*. But even among psychologist-authors, tradition dies hard, and such attempts to teach psychology in conformity with psychological principles still are rare enough to justify the reference (on the dust jacket) to Dr. Fleming's 'bold re-arrangement.' The main sections are: The Teacher as a Student of Motivation, as a Promoter of Learning, as an Observer of Growth, as Craftsman and Technician, as Experimenter, and as Administrator and Therapist. Unfortunately, the plan does not quite work out as it was conceived. For one thing, the allocation of many topics to their particular sections seems somewhat arbitrary. More seriously, the plan appears to break down within the sections and to revert to the traditional pattern. For example, although there is some reference to Understanding, yet the implications of this for Explaining are not worked out; Explanation, surely a major part of current educational practice, is not even listed in the index. Generally, one could have wished for the establishing of a more rigorous relationship between psychology and the actual procedures the teacher will use.

For this reason, among some others, the teacher-reader may find the book somewhat disappointing. Its compactness, admirable in itself, does not make for readability. Many of the points it does make will be almost meaningless (or worse, misunderstood) by the teacher whose previous knowledge of psychology is slight. Indeed, many teachers might complain that the title is a misnomer and that the book is not in fact about teaching at all—as they understand the term.

In view of the immense labour its preparation must have involved, one would like to think that the criticisms just levelled at it could be dismissed if, in fact, and despite the implications of its dedication (To All Ye Who Teach), the book was not intended for teachers. With an alternative title, synonymous with but more elegant than, "Psychology: A Source Book for Tutors in Training College and Departments of Education," it would fulfil a most valuable function. J. M. THYNE.

GROSS, NEAL, MASON, WARD S., and MCEACHERN, ALEXANDER W. (1958). *Explorations in Role Analysis*. New York: John Wiley and Sons; London: Chapman and Hall, pp. 379, 70s.

Workers in all fields involving analysis of social behaviour find it necessary to refer to the 'role' occupied by a member of a group or society. Loosely, this means status, together with expectations of behaviour associated with it. But, in summarising the literature of anthropology, sociology and psychology, where this term is used, the authors often find lack of precision and confusion of definitions between, and often within, these fields. They challenge particularly the explicit and implicit acceptance in many definitions of 'role consensus,' or agreement of expectations of behaviour in a given status among all group members, including the incumbent of the position; and this, together with other hypotheses, in their opinion requires investigation on an empirical basis. The second part of the book presents the results of an enquiry into this problem, and the last few chapters consider some psychological implications arising out of these.

To British readers, the specific subject of study, namely, the role of school superintendents in Massachusetts, may have less significance than the general methods of approach described. As a methodological guide the book should be very useful in suggesting what sort of information needs to be obtained, what questions should be asked and how to tabulate statistical information. A mass of data has been handled by up-to-date statistical methods, and the style of presentation is scholarly. There is a good summary of references at the end of each chapter, though the absence of a general bibliography is regretted.

In the last few chapters, 'role conflict' is used to describe a situation in which a group member occupies two roles simultaneously, or finds himself attempting to satisfy incompatible expectations from different sources in a single role. An example cited is that of the superintendent as a school official and a family man. The penultimate chapter attempts to analyse the evidence on the resolution of role conflict, but certain comparatively tenuous assumptions involved in the statement of the hypotheses tested make the conclusions rather unconvincing; for instance, on p. 302 the authors are obliged to admit the possibility of two theorectical models for their null hypothesis.

Social psychologists may feel that the methods used, particularly in the section on "Microscopic Consensus Analysis" in which small groups are involved, might have gone beyond the use of interviews and questionnaires to include a study of groups in action. While this particular study may not have lent itself to such methods, the work as a whole is so prolific in suggestions of lines of enquiry that it may well be the pioneer of a number of investigations using all available techniques.

D. M. EDWARDS PENFOLD.

HEIDER, F. (1958). *The Psychology of Interpersonal Relations.* New York: John Wiley; London: Chapman and Hall, pp. ix+322., 50s.

This is an attempt to treat interpersonal relations as distinct from the study of the individual or the group. Factor analysis may isolate distinct personality traits, but no person lives in a vacuum. There is the effect of p (person) on o (the other being), and the relationship between them and that of their environment. The author has evoked a whole notation for representing these relationships partly as he says fashioned after that of symbolic logic. He feels that a translation from the narrative of an event into notation can help to clarify the situation. Whether it does or not depends on the reader's own style of thought. Heider acknowledges a debt to Lewin and the topological school but finds their concepts inadequate for dealing with two persons situations. Indeed, Heider is nearer the academic psychology of Stout: "the awareness which we ordinarily have of ourselves is inseparably connected with the thought of other selves as variously related to us in the way of social intercourse," etc. It is the phenomena of the social situation which are considered here as desire and pleasure, sentiment, request and command, reaction to the lot of the other person. The author compares Piaget's work on the growth of concepts in the child with the change from mere perception to concept.

This is stiff reading, but it is a valuable study of the psychological factors in the relationships between individuals from an author whose bibliography does not include Freud but comprises philosophy and literature.

M. MARKHAM.

JOURARD, S. M. (1958). *Personal Adjustment.* New York: Macmillan, pp. xiii+462, 38s. 6d.

Recent works on counselling have shown a growing trend towards education in mental health for all rather than the former exclusive concern with remedial therapy for the deviate alone. This book marks a turning point in emphasis since, as its sub-title indicates, it approaches personal adjustment "through the study of the healthy personality." Born of a course given to American university students, and liberally supplied with chapter summaries and questions for discussion, it is intended primarily for use as a text in college courses on mental hygiene. Well documented, it makes wide use of recent American research in the psychology of personality growth and change, the main ideas deriving from Horney, Fromm and Maslow. With such a positive frame of reference, the author has found himself inevitably committed to a consideration of values. He believes that the healthy personality should reflect the best values for man that prevail in the society in question. Writing for Western democracies, his explicit aim is "the growth of the self towards self-actualization" and the book is permeated by an implicit philosophy of joyous abundance—"an economics of plenty rather than scarcity economics." The belief that "to him that hath shall be given" is particularly evident in the valuable chapters on relationship and love in which the author closely follows Fromm's *The Art of Loving*. The general approach is practical and sound, well suited to the level of student needs and problems. Deeper philosophical and spiritual concepts will be sought in vain, but once English readers have mastered the somewhat technical vocabulary, they will find the book simple, straightforward, realistic and useful for courses and discussion groups on personality in training colleges and universities.

MARY SWAINSON.

KAMAT, V. V. (3rd edit., 1958). *Measuring Intelligence of Indian Children.* Bombay: Oxford University Press, pp. xvi+263, 17s. 6d.

The growing number of books from Indian sources and the fact that this book is now in its third edition, emphasises the interest in and the demand for knowledge of test construction and the techniques of standardization. There is a very comprehensive account of a revision of the Stanford-Binet tests to suit Indian children. The test material was changed to suit the Indian background and great care was taken in assigning the tests to appropriate age levels. Administrative details for each sub-test are carefully examined and clearly expressed. From the practical point of view, this is a workmanlike attempt to provide a means of testing and assessing the intellectual ability of children which will provide an influential stimulus to other Indian psychologists.

Since this is the third edition of a book, first published in 1940, it is a pity that the opportunity was not taken to bring the contents of the introductory chapters up to date; less emphasis would then have been given to Educational and Achievement Quotients and to the cessation of growth of intelligence at 16 years.

W. H. KING.

LAFITTE, P. (1958). *Social Structure and Personality in the Factory.* London: Routledge and Kegan Paul, pp. xii+228, 24s.

The book represents the fourth volume in a series of publications from the Psychology Department, University of Melbourne. These are concerned with the understanding of man in his social context. While there is ample cross reference to the other volumes, Dr. Lafitte's approach is rather unique in methodology and analysis.

The material presented, is based on a large-scale investigation, carried out in Victoria, into the behaviour of persons in industry. It represents a welcome addition in a field which to some extent has been neglected by psychologists. Dr. Lafitte emphasizes the necessity for relating the behaviour of persons within industry to a wider social context. This is an important departure from the over simplification often apparent in texts in industrial psychology. Hence, this book is recommended, not only for those interested in industrial relations, but also as a useful general introduction to those who prepare others for entry into industry.

F. M. KATZ.

MACPHERSON, J. S., Scottish Council for Research in Education (1958). *Eleven-year-olds Grow Up.* University of London Press, pp. xix+184, 15s.

This report is a follow-up study of a sample of children already considered in the 1947 Survey. The sample consists of children born on the first day of each of the six months, February, April, June, August, October, and December, in 1936. This six-day sample consisted of 1,208 boys and girls and was 1·6 per cent. of the age group tested in 1947. No significant difference was found between the six-day sample and the children of the 1947 Survey in regard to size of family, geographical district, test scores and occupational class of father, and for this reason the six-day sample was taken as an unbiassed sample of the 11-year-old population in 1947. The age of eighteen years was taken as a suitable stage to report progress to date because at this age most of the adolescents had finished their secondary school courses and were adults in the eyes of the Ministry of Labour and National Service.

In addition to the test results provided by the 1947 Survey, information was now obtained from the schools, the first in 1950, when the children were 14 years old, and the second when the children left school. Home visitors and Youth Employment Officers obtained the necessary data after the children left the schools. The wealth of additional information classified under thirty-seven headings from these sources, makes a summary difficult. The main follow-up studies deal with reasons for lack of success in school courses, wastage among early leavers, happiness at school, effects of home background, changes of employment, leisure interests and attitude to National Service. All these are interesting and important, but because of the smallness of the numbers involved, only tentative conlusions are put forward. For example, an attempt is made to assess the suitability of jobs taken up after leaving school; it is deduced that "the more intelligent do not change as frequently as those of lower ability" and as regards study continued after leaving school, it is suggested that the classes for boys are designed to cover a longer period of time than those for girls, and especially for the boys, the employers' attitude is one of support for their continued training. In a study of the sociological factors which are important for educational success, overcrowding in the home and occupational class of father stand out. In its comments on 'wastage,' the report suggests the importance in selection for secondary education of the teachers' character ratings of their pupils and gives some evidence that ratings by teachers on conscientiousness, desire to excel and perseverance help to determine those children who are likely to be successful in a prolonged academic course.

Many of the points discussed in the body of the report are re-emphasised in the last chapter where comparison of the data obtained by high and low scorers on the Terman-Merrill are compared. The high scorers tended to be members of small families, ; more than half of the high scorers came from homes where fathers were in the upper range of occupational classes ; high scorers tended to come from homes where there was no overcrowding ; teachers' ratings of personality were related to I.Q. either because of a halo effect or because "high I.Q. and desirable personality characteristics generally go together." Unfortunately, the data were not able to give any clear indication of the influence of children's I.Q. on their teachers' assessments of character.

This report, like all the previous surveys by the Scottish Council for Research in Education, is thorough and well documented. Some psychologists may like to give their own interpretation of some of the eighty-six tables on which this report is based but few, perhaps, would express their results with as much restraint. The appendix includes the School and Home Visitors' Schedules and the Employment Record Card by means of which the follow-up material was collected.

W. H. KING.

PEERS, R. (1958). *Adult Education*. Routledge and Kegan Paul, pp. xiv+365, 35s.

The opening chapters of this book trace the history of Adult Education from its inception in the early nineteenth century to the complexity of its present-day form. Later chapters deal with the nature of students, tutors, courses and methods of teaching. The final chapters offer accounts of comparable educational facilities in several other countries. While the immediate nature of the book is neither psychological nor sociological, the subject-matter is constantly intertwined with social and political change, with the increased mobility of the working population, with the judgment of human interests and capacities, and with other such aspects of the broader educational pattern. Thus, the author provides much valuable material for research workers in a number of allied fields.

Of particular interest to psychologists are the sections on adult learning, tutors, and methods and of teaching. The former raises pertinent questions on the nature, resilience and depth of adult learning, setting these against the individual student background of age, previous education, and occupational experience. Existing work on adult intelligence and learning by Wecshler, Thorndike, Lorge, Welford, and others, is set against practical experience. The suggestion is made that formal psychological experiment in these fields takes no account of the combination of intelligence and maturity of experience which fosters adult learning, and that further research is needed here. It is surprising that Hebb's theory of intelligence does not find consideration.

Educational psychologists will note many similarities between school and adult education fields. The importance of interest, usefulness and achievement as incentives to effort and progress is seen against a wider and more varied background. The qualities of a good tutor correspond closely to those expected of a good teacher, and some correlate with those sought by training department tutors, as shown by Robertson in a previous issue of this *Journal*. Suggested methods of teaching have much in common with those fostered at the younger age-levels, indicating continuity in the educative process and in human relationships throughout.

Perhaps the most important single social thread emerging from the book is the constant struggle to learn more and more in the early years of life with the accompanying need to reconsider and rethink the continuity and enrichment of education in the later years. Here the reader is clearly shown the breadth of research and development required if society is to continue to enjoy a measure of stability and mental health.

D. M. Lee.

Raven, J. C. (1958). Revised editions : *Guide to Using the Coloured Progressive Matrices*, pp. 40, 5s. ; *Guide to Using the Mill Hill Vocabulary Scale with the Progressive Matrices Scales*, pp. 64, 5s. 6d. London : K. H. Lewis.

Almost all educational and clinical psychologists must have used Raven's Matrices and/or Vocabulary tests at some time ; but there are so many versions of each that it is useful to have these definitive manuals setting out the age and ability ranges, and the clinical purposes, for which each one is suitable ; (we are glad to learn that a similar revision of the standard Progressive Matrices Guide is also in preparation). While these do not incorporate any essential changes in the materials or manner of application of the tests, they contain considerably more information on interpretation of results, standards of performance over the 65-85 age range, and new norms for Sets A, Ab and B based on representative samples of Dumfries children.

The Vocabulary Guide includes an historical introduction and some discussion of the psychology of language development, but is noteworthy chiefly for the detailed consideration of abnormal responses to seventeen selected words (The Short Form). The largest numbers of such responses occur in hebephrenia, organic and senile psychoses. While it would be too much to claim that definite connections are established between different types of unusual response and different clinical categories, the classification should be of interest to users of any type of (creative-response) vocabulary test.

The Coloured Matrices A, Ab and B constitute an expansion of the standard Sets A and B, which are particularly suitable for children below 11 and old people. Again, a qualitative analysis of erroneous responses is provided, though the relation of these to different clinical conditions—if any—is not made clear. However, there are interesting patterns of relations between Matrices and Vocabulary in normal and maladjusted children, and in healthy, senile and depressive persons over 70.

P.E.V.

Reeves, Joan Wynn (1958). *Body and Mind in Western Thought.* Pelican, pp. 403, 5s.

This fat little Pelican book meets a long-felt need for a collection of writings on Mind and Body from earliest times to the present day. Dr. Reeves' anthology starts with the early Greeks, continues through St. Augustine, Francis Bacon, and, of course, Descartes, Spinoza and Leibniz, and on to Wundt and James. It is set out in two parts : an excellent historical introduction, followed by extracts from the writers themselves with editorial comment. It will not only be very useful as a book of reference with many quotable passages, but also provides a fascinating account of the development of thought and opinion about the mind/body relationship. For those interested in these problems it is five shillings very well worth spent.

R. Hetherington.

Semeonoff, B., and Trist, E. (1958). *Diagnostic Performance Tests.* London : Tavistock Publications, pp. xv+176, 32s.

This manual describes adaptations of some individual exploratory procedures : the originals were the Vigotsky, Weigl, Kohs, Carl Hollow Square, and Passalong tests. Half the book is devoted to norms, statistical analyses, and appendices. Most of the work was done about twelve years ago.

These procedures study the detailed cerebration of the individual, how he perceives, categories, surmounts difficulties, perseveres, perseverates, or peters out: all as practically important as any global intelligence index. This volume is recommended to those who wish to extend their testing skills in this direction; these adaptations are sensible, and the materials can be bought easily. Some parts could, with discrimination, be used with children: there is a plain opportunity to devise an analogous series, particularly for handicapped children.

The dilemma with this type of test is whether one gets more by observational recording or by the more cramping conditions necessary if standard norms are used. The authors' own approach is not rigid, but the elaborate marking systems may tempt some to concentrate on summed scores rather than study individual problem-solving patterns, wherein the value of these methods lies. (Ultimately, perhaps, test-norms will be used as an incidental sighting-point, not to give a final 'social placing' (like an I.Q.); as in physical medicine tables of normal heights and weights are for initial orientation, not for diagnosis). More of the book (overloaded with minutiae) might have been given to discussion of the psychological concepts involved.

Professor Vernon's foreword notes some professional difficulties in developing this field, particularly the 'brand-image' of the psychologist as an I.Q. technician.

Layout and photographs make the price high. Burt the book is excellently produced; though it loses a point in the ancient battle to get printers to set 'noegenetic' correctly. It will be a standard British reference manual for such tests for some time to come.

B. CLARKE.

SOFER, C., and HUTTON, G. (1958). *New Ways in Management Training.* London: Tavistock Publications, pp. xii+127, 15s.

This is a study by the Tavistock Institute on the Department of Management and Production Engineering at Acton Technical College, London. It was sponsored by the Department of Scientific and Industrial Research and the Medical Research Council.

This report might well represent a turning point in the role played by the technical college in relation to industry and will stimulate some re-thinking about teaching methods in advanced courses. As its publication had to receive the consent of interested parties it is naturally couched in cautious, diplomatic language, and under these circumstances the authors have done well in retaining the interest of the reader and including a wealth of constructive proposals. Though none of these can be said to be new, in total they represent a re-orientation of outlook. Closer links between the Department and industry are advocated, particularly the formation of 'Developmental Panels,' more lecturing by 'visiting consultants' from industry, and the carrying out by teachers and students, of projects within industry. To this end, it points to the rigidity of the regulations governing the duties of the staff as one of the principal barriers preventing an increase in effectiveness. During the survey a promising beginning was made in student counselling.

Of value to psychologists in industry and to educational psychologists who are interested in technical education, the survey will also serve as a model of the manner in which field work may be carried on without causing a disastrous treading on tender toes. This success was helped by the initiative coming from the Department, but key methods were the 'incorporation' of the staff in the research, and the presentation of the results as constructive proposals.

R. R. DALE.

WINTERBOURN, R. (1958). *Caring for Intellectually Handicapped Children.* Wellington: New Zealand Council for Educational Research; London: Oxford University Press, pp. viii+47, 4s. 6d.

SCHONELL, F. J., RICHARDSON, J. A., and MCCONNEL, T. S. (1958). *The Subnormal Child at Home.* London: Macmillan, pp. vi+167, 5s.

Professors Winterbourn's and Schonell's little books both consist of advice to parents of feeble-minded children, and are similar to our own National Association for Mental Health's Handbook, *Mentally Handicapped Children*, and to several American publications. The differences between the welfare and educational services in New Zealand, Australia, and Britain, are not great enough to make them inapplicable here. Sound and sympathetic suggestions are given on the family problems that arise, whether to commit the child to an institution, how to train him, where to turn for further help, and many other difficulties.

Professor Schonell's is the more detailed, but suffers to some extent in that it refers to high-grade as well as low-grade children ; thus, it is rather frustrating, perhaps, for parents of an idiot to be told about elementary reading and number work and vocational adjustment. On the other hand, it contains much that will be helpful to teachers in E.S.N. schools. But it is surprising, to say the least of it, to find the statement that : " research carried out in . . . England on mental retardation is practically negligible." Professor Winterbourn's pamphlet, which is being distributed free to all parents of handicapped children in New Zealand, is definitely confined to the lowest 0·5 per cent. of the population.

PUBLICATIONS RECEIVED

The mention of a book in this list neither implies, nor precludes, a later review.

A. M. Blackwell (1958). *Lists of Researches in Education and Educational Psychology, Supplement III.* London : Newnes, pp. 64, 15s.

M. Sakel (1958). *Schizophrenia.* New York : Philosophical Library, pp. ix+334, \$5.00.

M. Sakel (1958). *Epilepsy.* New York : Philosophical Library, pp. xxxvi+204, \$5.00.

J. R. Gallagher and H. I. Harris (1958). *Emotional Problems of Adolescents.* New York : Oxford University Press, pp. xiii+174, 28s.

H. E. Garrett (1958). *Statistics in Psychology and Education*, 5th ed. New York : Longmans Green, pp. xii+478, 35s.

H. T. Himmelweit, A. N. Oppenheim and P. Vince (1958). *Television and the Child.* Oxford Uuiversity Press, pp. xix+522, 42s.

I. Vikainen (1958). *Synopsie und ihre Bedeutung fur die Didaktik.* Turku, Finland : Turun Yliopiston Kustantama, pp. 216, 800 Marks.

V. L. Senders (1958). *Measurement and Statistics.* New York : Oxford University Press, pp. xvi+594, 48s.

N. M. Downie (1958). *Fundamentals of Measurement : Techniques and Practices.* New York : Oxford University Press, pp. xi+413, 48s.

P. Halmos and A. Iliffe (ed.) (1959). *Readings in General Psychology.* London : Routledge and Kegan Paul, pp. ix+251, 25s.

J. Piaget (1959). *The Language and Thought of the Child* (3rd edit. rev.). London : Routledge and Kegan Paul, pp. xxiv+288, 25s.

A. Adler (1958). *The Education of the Individual.* New York : Philosophical Library, pp. xiii+143, \$3.50.

G. F. Kneller (1958). *Existentialism and Education.* New York : Philosophical Library, pp. xiii+170, \$3.75.

R. M. W. Travers (1958), *An Introduction to Educational Research.* New York : Macmillan, pp. xviii+466, 40s.

C. C. L. Gregory and A. Kohsen (1959). *The O-Structure: An Introduction to Psychophysical Cosmology*. Church Crookham: Institute for the Study of Mental Images, pp. 141, 21s.

H. M. Lynd (1958). *On Shame and the Search for Identity*. London: Routledge and Kegan Paul, pp. 318, 25s.

N. L. Goddard (1958). *Reading in the Modern Infants' School*. University of London Press, pp. 142, 13s. 6d.

R. Hutton (1958). *Of Those Alone*. London: Sidgwick and Jackson, pp. 235, 18s.

M. Balint (1959). *Thrills and Regressions*. International Psycho-Analytical Library, No. 54. London: Hogarth Press, pp. 148, 18s.

P. Fletcher (1958). *Emotional Conflict*. London: Duckworth, pp. 143, 10s. 6d.

World Federation for Mental Health (1958). *Growing Up in a Changing World*. London: W.F.M.H., pp. 238, 15s.

K. M. Miller and M. A. Nicol (1958). *Occupational Status, Sex and Age as Factors in Radio Programme Choice*. University of Tasmania Department of Psychology, Publication No. 1. Hobart, Tasmania, ; pp. 79, no price given.

H. A. Goddard (1958). *Principles of Administration Applied to Nursing Service*. World Health Organization Monograph Series, No. 41. Geneva: W.H.O., pp. 106, 20s.

Annual Review of Psychology, Vol 10 (1959). Palo Alto: Annual Reviews Inc., pp. x+520, $7.00.

A. R. Holt and L. Luborsky (1958). *Personality Patterns of Psychiatrists*. Menninger Clinic Monogr. Ser. No. 13. New York: Basic Books, Vol. I, pp. xiv+386, $7.50; Vol. II, pp. xiv+400, $4.00.

VOLUME XXIX. NOVEMBER, 1959 PART III

A FOLLOW-UP STUDY OF SOME FORMER GRADUATE STUDENT TEACHERS

BY MILDRED COLLINS

(*Department of Education, University of Leicester.*)

SUMMARY. A follow-up study was made of 115 former graduate student teachers of one education department and assessments by heads of schools were compared with teaching grades awarded during training. There was a 75 per cent. correct forecast as to which students were likely to have, or not to have, difficulties and a biserial correlation of +0·566 was obtained between heads' and tutors' assessments.

A study of discrepancies in assessments (' over ' or ' under ' achievement) was used as a check on conclusions derived from an enquiry concerned with teaching practice assessments only. ' Under ' or ' over ' achievement in schools seemed to indicate the importance of type of school, academic achievement, physical and mental health, family problems and length of professional experience among former students who had shown rather limited practice teaching ability. Case histories demonstrated how difficult it is in individual cases to distinguish the ' late-developer ' from the ' non-starter.'

I.—INTRODUCTION.

VERY little is known in this country about the relation between teaching assessments of students during training and their teaching in schools afterwards. In 1942, Tudhope[1] reported on the careers of former training college students and found correlations of +0·84 for men and +0·77 for women between practice teaching assessments and assessments by H.M.Is. in schools afterwards; but much more recently Nisbet[2] found no association between Department teaching grades of a Scottish university and status of teaching posts held by former students 20 years later. Some discrepancies between practice and professional teaching achievements would seem inevitable, but marked discrepancies would bring into question the whole system of teaching assessment and the relation between the aims and methods of schools and those of training colleges and departments of education. In fact, in America, where correlation seems to have ranged between +0·16[3] and 0·348[4], one investigator[5] has gone so far as to question whether " practice and actual teaching are comparable difficulties." The rarity of follow-up studies here makes it impossible even to pose the question.

Obviously, such a state of affairs is undesirable, especially as research into the attributes of teachers is frequently based on practice teaching grades only. Conclusions derived from such researches are often conflicting, for example those of Phillips[6] and Lovell[7] on personality characteristics, of Lawton[8] and Evans[9] on extra-curricular activities and of Weinstock and Wood[10] and Burroughs[11] on the value of interviews. Certainly, some of these discrepancies can be explained by differences in selection of the student groups, in the type of questions asked, and the analysis of the replies received, but they do not add, at least in their present form, to a feeling of confidence in the reliability and validity of practice teaching assessments. Hence, when some research was carried out recently, into various correlates of practice teaching ability within an education department of a university,[12] it was considered wise to make a follow-up study of some of the students concerned, to see how far conclusions based on teaching practice grades still held when former students took up teaching posts in schools. This article will give an account of the follow-up study.

II.—THE PROCEDURE.

(i) *The Sample.*—That part of the research based on teaching practice grades was largely concerned with two groups of students drawn from six sessions. One group (the P. group) consisted of all students in the six sessions (with the exception of three) who had obtained practice teaching grades of C− and below ; these were the below-average students. The other group (the G. group) consisted of students ' matched ' with the P. group for degree standard, teaching subject(s) and hence method tutor, type of teaching practice school, session, sex and age, but with a teaching grade of C or above ; these were the students of average, or above average, teaching ability.

There were eighty-four pairs of students in the original groups, but it proved impossible to obtain information about all the members, and seven P. and eight G. students were still doing National Service and so could not be included. In all up-to-date information was obtained for fifty-eight P. and fifty-seven G. students, and it is on these that the follow-up study is based. The distribution of the teaching grades of these students may be seen from Table 1 ; fortunately, the proportion of the different levels of teaching ability represented do not differ significantly from those present in the original teaching practice groups.

TABLE 1

TEACHING PRACTICE GRADES OF STUDENTS.

Groups of Students	P Students			G Students			Total
	E	D or D+	C—	C or C+	B—or B	B+ or A—	
Teaching Practice	5	31	48	43	32	9	168
Follow-Up Study	4	18	36	26	25	6	115

($\chi^2=2{\cdot}537$, d.f.$=5$, $P={\cdot}80<>{\cdot}70$).

The proportion of women among the follow-up and teaching practice groups was virtually the same, 29 per cent. and 30 per cent., respectively ; among the students as a whole women made up 33 per cent. The schools in which the students took teaching posts are shown in Table 2 ; the significant

TABLE 2

TYPES OF SCHOOLS FOR ASSESSMENT OF TEACHING.

Occasion of Assessment	Grammar	Secondary Modern	Junior	' Other Schools '	Total
Teaching Practice	84	54	26	4	168
Permanent Teaching ..	44	29	15	13	101*
Expected Frequency of Posts in Schools	50·4	32·4	15·6	2·4	101

$\chi^2=48{\cdot}0$, d.f.$=3$, $P=<{\cdot}001$.

* Excludes, of course, the fourteen students who did not take posts in schools.

change was a reduction in grammar and modern school posts and an increase in those in ' other schools ' ; that is in technical schools and colleges and private schools. Fourteen of the 115 students did not take up teaching at all, although qualified to do so. We may conclude that the follow-up group is a fairly representative sample of the students concerned in the original enquiry.

(ii) *Assessment of Pre-service and In-service Teaching.*—Pre-service, or teaching practice, assessments were carried out by the method tutors of the students in consultation with other tutors, the professor of education and, in some cases, with H.M.Is. or tutors from other education departments. It was essentially a department decision built up over the term's teaching experience but modified in some cases by an outside opinion on a particular lesson.

In the follow-up study the in-service assessments were based solely on the opinion of the head of the school in which the former student was teaching or had taught. This seemed the only way of assessing students scattered widely throughout the country, and a head's assessment is obviously a most important one for the teacher concerned. It may be argued that headmasters and headmistresses have individual standards influenced by the characteristics of their schools and the Staff they have become accustomed to appoint. But these difficulties confront Department assessors also to some extent, and what little evidence emerged from this enquiry concerning a common standard for head's assessments was reassuring. It so happened that for eight students (which included poor, average and good) there were assessments from two or more heads and in no case was there a reversal of judgment, although with two science graduates the assessments—not unexpectedly—became more favourable as time went on.

The periods of teaching experience on which the assessments were made varied of course with the age and year of training of the student and ranged from one to seven years, the most commonly occurring periods being one and four years (with thirty-seven and twenty students, respectively). The number of assessments based on one year's experience was boosted by those students who were unsatisfactory and left a school after one year's teaching and also by those who left to get married. A resumé of the letter which was sent to the head of the school at which the former student was thought to be teaching is given in the Appendix ; 82 per cent. of the headmasters and headmistresses replied and gave assessments. These assessments were made as simple as possible and three categories only were employed " Less well than usual," " Quite well " and " Very well." A space was left for comments and with few exceptions heads took advantage of this ; some sent explanatory letters also.

The assessments from the schools were then analysed to see how far they coincided with those made in the education departments, and those discrepancies which occurred were considered in the light of the conditions known to have been associated with achievement during the teaching practice period.

III.—THE RESULTS OF THE FOLLOW-UP ENQUIRY.

(i) *The relation between practice and professional teaching assessments.*—A statistical analysis of the two teaching assessments was made and also a review of ancillary evidence concerning the relative achievement in schools of the P. and G. groups. The analysis was complicated in this enquiry, first by the method of selection of the groups which resulted in the P. group including 67 per cent. of the total number of below average students while the G. group included only 14 per cent. of the average or above average students in the department. Secondly, there was a lack of normality in the distribution of the assessments from the schools. A further difficulty, likely to be encountered in all enquiries of this kind, was what to do with those students who, though qualified to do so, did not teach in schools ; there was no means of knowing how those students would have fared in schools. As no single approach seemed satisfactory, the analysis was carried out in three ways.

First, there was a simple four-category assessment ; students were grouped according to whether they had never taught or had left teaching, were doing " Less well," " Quite well " or " Very well." (Women students who had left teaching after marriage were classified according to their teaching in schools and were not put in the " left teaching " groups.) The analysis was carried out for the P. and G. groups separately, as can be seen from Table 3, and the Chi-square test applied to test the hypothesis that there was no significant difference between those former students classified as below average, and those classified as average or above. Actually $\chi^2=22{\cdot}10$, d.f.$=3$, $P=\cdot 001$; this shows a highly significant difference remaining between the two groups in spite of fourteen rather poor students having developed into good teachers.

TABLE 3

PRE-SERVICE AND IN-SERVICE TEACHING ACHIEVEMENTS.

Pre-Service Groups	In-Service Teaching Achievements					Total
	None at Enquiry		Heads' Assessments			
	Never taught	Left teaching	Less Well	Quite Well	Very Well	
'P' Group	9	9	14	12	14	58
'G' Group	5	1	2	21	28	57

'Never taught' and 'Left teaching' were grouped together in calculating Chi-squared.

$\chi^2=22{\cdot}10$, d.f.$=3$, $P=\cdot 001$.

Secondly, an attempt was made to get an estimate of the degree of association between teaching under the two conditions by calculating the biserial correlation between the two sets of grades, using a finer division of categories for the school assessments and converting them into scores. The finer divisions were obtained by using the comments added by the heads to the single three-category assessment, so that each original category became divided into three, with the lowest score in each indicating an 'unfavourable' comment and the highest indicating a 'favourable' comment within that category. In addition, those students who had never taught were given a score of 1, and those who had left teaching a score of 2. The scores within the assessment categories then continued from 3 to 11. The mean score for the P. group was 4·702 and for the G. group 7·709 ; the S.D. of both groups together was 3·301.

The scores were correlated with below average and average and above teaching ability during training and the biserial r was 0·566.* This is significant but falls much below the unexpectedly high correlation obtained by Tudhope.[1]

---

* After adjustment for differing size of samples included in the P. and G. groups, biserial $r=+0{\cdot}339$.

A third method of assessing the relationship between school and department gradings was to consider the percentage of errors made, that is the proportion of below average student teachers who proved satisfactory in permanent posts or who proved so inadequate that they had to leave—and the proportion of the better students who were unsatisfactory in the permanent school situation. In this method students who had never taught were excluded from consideration. Presumably, all the below average student teachers were expected to have some difficulties at first, but presumably also they were expected to become adequate teachers in time. In estimating the number of errors a score of 6 was taken as the dividing line for the school assessments. This was obtained by a student doing "Quite well" but who had had difficulties sufficiently obvious for the head to comment on. Any below average student scoring above 6 (i.e., no difficulties commented on) was counted as an 'error' and there were sixteen of these students, while any average or above average student with a score of 6 or below was likewise counted as an error and there were four of these. In addition, there seems a case for including as errors those P. group students who had to leave teaching because of difficulties; presumably these students should have been failed in the first place and there were nine of these.

Using these rather stringent criteria we have twenty-nine errors in all among 101 students, giving a proportion of 29 per cent. If we are less stringent and exclude those students who have left teaching but who were recognised as being very 'doubtful starters' during training and were given a D teaching mark, then we remove four 'errors' and we have 25 per cent. wrong assessments for these particular groups. One has to remember, however, that the G. group represents a much smaller proportion of better students than does the P. group of the poorer students, and it is the former, while using the criterion of early difficulties in teaching posts, which has the fewer errors. However, no matter by which method the percentage error is calculated, it in no way supports Bach's[5] suggestion that practice and actual teaching are not comparable situations: in this enquiry the poorer practice teachers tended to become the poorer actual teachers.

Ancillary evidence concerning the careers of the members of the P. and G. groups supports the evidence from the head's assessments, although the numbers concerned are small. Ten of the G. group but only two of the P. group were recorded as having obtained graded or head of department posts and in both cases the P. students were from early sessions and the posts were in schools different from those in which they did their teaching practice. This difference in status of teaching post is contrary to Nisbet's[2] finding, with teachers of much longer experience.

Perhaps unexpectedly five out of the fifty-seven G. students did not take up teaching, but there were almost twice as many among the P. students who did not do so either. A minor but interesting difference between the two groups is shown in the number of women leaving teaching on marriage. Information was available for five of the P. group and seven of the G. group who married as teachers; only one of the former taught for twelve months after marriage but all seven of the others did so. It seems that those students who were free from noticeable difficulties during training, i.e., those of the G. group, were, as one would expect, keener to teach and more successful in acquiring responsible posts.

(ii) *The discrepancies between the two assessments.*—Nevertheless, in spite of the positive correlation between practice and professional teaching, there were in this enquiry some striking differences between achievement in the two situations, and it is important to consider these more closely.

That the time lapse between assessment in training and the assessment in schools influences the apparent validity of the former has been suggested by the award of graded posts to P. students from the first two sessions only. Students from the first session could have taught for seven years when the enquiry was made while those from the last session could only have taught for a year. Hampton[13] showed by an analysis of supervisory ratings in the States that—not unexpectedly—discipline improved over a two-year period. The importance of time for the P. students of this enquiry can be seen by looking at Table 5 which shows a higher proportion of students in the " Very well " category for the first three sessions compared with the last three. Time has presumably remedied and erased from memory their former weaknesses.

TABLE 5

LENGTH OF TEACHING EXPERIENCE AND TEACHING ABILITY IN SCHOOLS.

Sessions	Number of Students		Number of Students assessed as ' Very Well '	
	P. Group	G. Group	P. Group	G. Group
1, 2 and 3........	24	27	9 (37·5%)	13 (48·1%)
4, 5 and 6........	34	30	5 (14·7%)	15 (50%)

However, not all the poorer teachers improved with time, fifteen former students (including two from the G. group with C, C+ teaching grades) have either left teaching, been advised to do so, or had their probationary period extended. These students have done worse than expected and will be called the ' under-achievers.' On the other hand, sixteen of the P. group had no recorded difficulties such as one would expect : these we shall consider the ' over-achievers.'* It is proposed to test out some of the conclusions based on teaching practice assessments only on these two groups, the hypothesis being that the over-achievers will show more of the characteristics positively associated with teaching practice success than will the under-achievers. But before doing this it should be noted that some of the differences between over and under achievement can be accounted for in terms of length of experience in schools :

TABLE 6

TEACHING PRACTICE GRADES OF OVER AND UNDER-ACHIEVERS.

	Students with T.Gs. of :			
' Sub ' Groups	E, D or D+	C—	C or C+	Total
Over-Achs :	3	13	—	16
Under-Achs :	7	6	2	15†

---

* We shall ignore, for the purposes of this enquiry, those average teaching practice students who did really well in schools and also those outstanding students who did not quite live up to their early promise.

† Includes 2G. group students.

57 per cent. of the over-achievers compared with 37 per cent. of the under-achievers came from the first three sessions. Some difference in achievement was, however, present even before the two groups left the Department: Table 6 shows that the better quality of the over-achievers was realised by the Tutors although not to the extent that it materialised later.

However, to turn to the verification of conclusions based on teaching practice assessments[12] only: it was predicted that the over-achievers would:

(*a*) Have better academic qualifications.
(*b*) Have transferred to a different type of school (bearing in mind the importance of the school and their comparative failure in the teaching practice schools).
(*c*) Be in better health and have fewer personal problems.

Reference to Table 7 shows that, in fact, the over-achievers were better academically, both in degree work and in that for the Certificate of Education; and change of school was more frequent among the successful students. Possibly this apparent advantage derives not only from different teaching demands in a new type of school, but also from a more flexible attitude to problems in general which this change implies. In connection with a possible difference in attitude, it may be important that three of the over-achievers had taught abroad for a short period, but none of the under-achievers had ventured so far.

TABLE 7

CHARACTERISTICS OF OVER AND UNDER-ACHIEVERS.

2 Sub-Groups	Characteristics				
	Hons. Degrees I or II	'Goods' in Cert. Subjs	Diff. Schools	Health Good	No Problems
16 Over-Achs:	9	13	9	6	7
15 Under-Achs:	3	8	4	9	8

Information regarding the health and personal problems of students was obtained from the health records of the department, attendance at the psychological service, some special interviews, and news of students since leaving. It is not likely to be complete, particularly for the practising teachers. The comparison between health and problems of under and over-achievers during training suggests at first glance that the unhealthier the student and the more numerous his problems, the better the prognosis from the point of view of teaching! In one sense this is true, worry and illness during training may explain a temporary poor achievement: seven of the over-achievers are known to have had health or family problems which have cleared up since the students have left the department, while among the under-achievers there are three cases known where problems have arisen since the students left. Some problems are likely to be more recalcitrant than others: five of the under-achievers but none of the over-achievers complained, while students, of family quarrels and jealousies. Possibly these family stresses of long standing generate or derive from unhelpful attitudes to life which become particularly conspicuous in teaching, for teaching in some ways reproduces a family situation.

Hence, it seems that the later success or failure of the poorer student teachers can be explained, at least in part, in terms of circumstance and personality rather than entirely in terms of wrong assessments during teaching practice or in schools. It would appear possible that over-achievers were persons ' who improve upon acquaintance ' or were ' slow starters.' This is suggested by the fact that while testimonials and interviews appeared to have about equal validity for the students as a whole during teaching practice, for the under and over-achievers the testimonials were the better indicators of teaching potentiality in the long run ; almost half of these students were correctly ' identified ' by the testimonials and only about one-eighth by the interviews. However, one must bear in mind, for all the conclusions based on under and over-achievers, that the numbers concerned were very small.

Nevertheless, the results point to conclusions which seem reasonable. The apparent errors of assessment stand out most clearly in the case of three students with teaching grades of D+ and below who did well in schools and of five students with teaching grades of C – (one C+) who did badly later and left teaching.

The three students who had improved so much by the time the follow-up study was made came from the earlier sessions and had changed both their type of school and teaching subject on taking permanent posts. Two of the three had illnesses of a psychological nature during the training year but have now recovered. The third student left the department without completing his teaching practice because he felt that he could not face classes. This sudden withdrawal was really repeating a family and personal pattern. After leaving, he took a non-teaching post, then was persuaded to do supply teaching and finally returned to the department at his own request to complete the course which he did quite successfully.* All these students have general degrees but only one teaches in a grammar school. Nothing is known concerning their family circumstances since leaving the department ; one certainly had a problem while within it.

Of the five ' failures ' only one changed the type of school on getting a permanent post and she continued to teach the same subject. Very little is known about one student so he will have to be excluded from further comment: during training it was remarked that his teaching was disappointing in the light of his activities within the student group. To take the four remaining students in turn :

(1) This student had a good degree and worked hard both at the university and in his school, but relationships with people were difficult. He had been given psychological help during training, but he had a family problem of very long standing : how far this situation has changed is not known.

(2) The second was a lively student with a poor Arts degree in spite of having a very high score on the Wechsler Bellevue Verbal Intelligence Scale. He had originally wanted to take up another profession, but his failure in teaching was a continuation of a lack of fulfilment of promise, against a background of family problems, which dated from his own school days.

(3) This student on the other hand,, was a quiet, older man with considerable academic ability, although he was not teaching in a grammar school. He had a pleasant personality and appearance, but was quite unable, according to the headmaster's report to make anyone listen to him. He had had various kinds of jobs both before and after taking his degree. His health appeared to be good and so far as is known he had no family problems, but he was a shy person whom a thoughtful headmaster found himself unable to help.

---

* When the original teaching-practice enquiry was taking place, this student was in his non-teaching post, but by the time the follow-up study was made he had returned to the Department, completed the course, and obtained a post.

(4) The fourth student also gave the impression that teaching was not her first love and implied that strong family persuasion and expectation had influenced her choice, but beyond this there seemed no family problems during training. However, one cropped up when she took a permanent post and this probably tilted the balance in favour of leaving the profession, although her headmaster, while admitting the difficulties, was prepared to ' battle on.' The student has a degree in a subject which makes it easy to find employment in industry. She is a pleasant, perhaps slightly subdued person, who enjoys good health.

The careers of these former students suggest two causes of failure, possibly inter-related. Cases 1 and 2 give the impression of failure in teaching as but one aspect of a failure to adjust to situations and people which has been in evidence for some years, while Nos. 3 and 4 point to a half-hearted commitment to teaching as a possible correlate of failure in border-line cases. But these instances, and those of the unexpectedly successful former students, indicate how difficult it is, on the information normally available, to decide who will survive and who will fail among students who begin with limited teaching ability.

IV.—GENERAL CONCLUSIONS.

It seems clear that, for the graduate teachers of this enquiry, there was a relation between teaching ability during training and in schools after qualification ; 75 per cent. achieved approximately what one would have expected from their teaching practice performance. The better students obtained more graded posts and posts of special responsibility ; the better women students appeared to stay longer in teaching after marriage. The association between practice and actual teaching attainment was not as close as that found by Tudhope for training college students, but his school assessments were obtained from H.M.Is. and these would tend perhaps to be more comparable to those of tutors than the assessments of heads of schools, who often take into consideration social acceptability in the staff common-room, etc. Department assessments were based on class-room teaching achievement only.

However, these differences in the basis of assessment would account in only a minor way for discrepancies ; for one thing a feeling of general acceptance by one's colleagues would tend to help one to teach rather better. A follow-up study of those former students who did less well or who did better than expected seems to show that the choice of school is an important factor. With poorer student teachers a very careful consideration of alternate forms of teaching seems desirable before sending them into permanent posts.

Time is on the side of those poorer students who manage to start fairly well in their first post. Not only does it naturally enable them to develop more skill, but it also provides an opportunity for health to improve or family problems to disappear—and these during teaching practice have been shown in this enquiry to be important. Also there is a slight indication that some poorer students are the kind of people ' who improve upon acquaintance,' and these in a permanent teaching post may come more into their own. The converse of, course, may hold true both for health problems and personality. It would seem that, perhaps, more care should be given to the physical and mental health of both practice and professional teachers.

It seems inevitable, if the practice teaching grade is to represent teaching attainment at that stage, that there should be discrepancies between pre-and in-service assessments. In fact, from the point of view of the schools and the poorer teacher they are to be welcomed, both by the elimination of the incompetent and the further development of those with greater potentialities. This is not to say that it would not be preferable if more of the former could be identified earlier. Research should be directed towards distinguishing the ' late-developer ' from the ' non-starter ' among those of rather doubtful ability. Evidence should also be collected on the teaching abilities required for different types of schools.

The results of this enquiry suggest that with suitable safeguards research into teaching can reasonably be based on work during the training period when other facilities are not available. They also indicate that follow-up studies, while important in their own right, may be useful as one means of checking conclusions based on training assessments only.

ACKNOWLEDGEMENTS.—I should like to say how grateful I am to those heads of schools who replied to my inquiry and thus made this study possible, and to those former students not teaching in schools who provided information also.

V.—APPENDIX.

Teaching Assessment Form sent to heads of schools.

" CONFIDENTIAL.

Please return to Mrs. M. Collins.

.... has now been teaching in this School since .... holding the post of .... and in comparison with other teachers of similar age and experience, has done

(*a*) very well.
(*b*) quite well
(*c*) rather less well than usual.

Any comments which you may wish to make to clarify assessment under (*a*), (*b*) or (*c*).

Signed."

A letter accompanied this form and stated that an attempt was being made to assess the predictive value of teaching practice grades of a sample of students, but no indication was given of the nature of the sample or that it could be divided into two groups. It was suggested that, if the head wished, the student might be told that the enquiry had been made after the completion of the form ; the head was also asked not to consult correspondence concerning the student's appointment until he (or she) had filled in his (or her) assessment.

VI.—REFERENCES.

[1] TUDHOPE, W. B. (1942). A study of training college final teaching marks as a criterion of future success in the teaching profession. *Brit. J. Educ. Psychol.*, 12, 167-171.

[2] NISBET, J. D. (1954). A twenty-year follow-up of intelligence test scores. *The Advancement of Science*, 11, 13-16.

[3] MARTIN, L. O. (1944). *The Prediction of Success for Students in Teacher Education.* Bureau of Publication, Teachers' College Columbia University, New York.

[4] SANDIFORD, C. W. and TRUMP, J. E. (1950). Teacher education, IV—Pre-service selection. *Encyclopaedia of Educational Research*. New York, : Macmillan, 1310-1396.

[5] BACH, J. O. (1952). Practice teaching success in relation to other measures of teaching ability. *J. of Exp. Educ.*, 21, 57-80.

[6] PHILLIPS, A. S. (1953). *An Examination of the Methods of Selecting Training College Students*. M.A. Thesis, University of London Library.

[7] LOVELL, K. (1951). *An Investigation into Factors Underlying Teaching Ability in Primary and Secondary Schools with a View to Improving the Methods of Selecting Potential Teachers*. M.A. Thesis, University of London Library.

[8] LAWTON, J. (1939). A study of some factors useful in choosing candidates for the teaching profession. *Brit. J. of Educ. Psychol.*, 9, 131-144.

[9] EVANS, K. M. (1952). *A Study of Teaching Ability at the Training College Stage in Relation to Personality and Attitudes of the Students.* Ph.D. Thesis, University of London Library.

[10] WEINSTOCK, M. B., and WOOD, M. D. (1951). An experiment in student selection. *Bulletin of Education*, 26, 14-18.

[11] BURROUGHS, G. E. R. (1958). A study of the interview in the selection of the students for teacher training. *Brit. J. Educ. Psychol.*, 28, 37-46.

[12] COLLINS, M. (1958). *Some Correlates of Inadequate Teaching Ability.* Ph.D. Thesis University of London Library.

[13] HAMPTON, N. D. (1951). The analysis of supervisory ratings of elementary teachers graduating from Iowa State Teachers' College. *J. of Exp. Educ.*, 20, 179-215.

MALADJUSTED CHILDREN IN GRAMMAR SCHOOLS

BY MAURICE CHAZAN

(*Educational Psychologist, Liverpool Education Committee Child Guidance Centre.*)

SUMMARY. This paper discusses the problems presented by sixty grammar school children referred to a child guidance centre between 1949 and 1956. Some of the children had difficulty in maintaining the standards of behaviour and work required of them ; some showed nervous or neurotic symptoms. The analysis of their case-histories suggests that in most cases their maladjustment was related to a combination of factors, but neither physical condition nor lack of ability was a frequent cause. There were severe personality defects or disturbances in only a quarter of the cases. In nearly every case there were adverse environmental factors. Faulty parental attitudes were particularly damaging, though often modifiable.

The methods of approach to treating the problems, and the outcome of treatment, are outlined. It is suggested that there is a need for further research in this field, and for the wider provision of special schools specifically catering for maladjusted children of superior intelligence.

I.—INTRODUCTION.

THE number of grammar school children seen at child guidance clinics is not large. Most grammar school children are able to meet the demands made on them without undue strain. There is no evidence to support the opinion that intellectually able children are usually physically weak or mentally unstable, though it should be recognized that they may have specific difficulties of adjustment (Wall, 1948). Terman (1926), in his study of gifted children, found that they were above average in physical growth as well as in general health. Mannheim (1948) mentions that he found the proportion of delinquents from grammar schools to be very small.

Between 1949 and 1956, sixty grammar school children (thirty-four boys and twenty-six girls) were referred to the Liverpool Education Committee Child Guidance Centre. In this study, their case-histories have been examined with a view to delineating the problems presented and eliciting the main causal factors. There is a varied school background, the children coming from nineteen different grammar schools. In nearly every case, information on the children was available from the head of the school and the home was visited by a social worker ; at the Centre, there were assessments by both a psychiatrist and an educational psychologist.

II.—PREVIOUS WORK.

Although there is now a vast literature on the general and individual needs of the dull and backward, little work has been done on the emotional difficulties of the child of superior intelligence. Schonell (1952) mentions the need for work in this field, which, he states, has received scanty attention from investigators. McClelland (1942), in his studies of misfits at the secondary school stage in Scotland, showed the importance of factors other than tests and gradings, in particular home conditions, on some children's school careers. Burns (1949) made a clinical study of eighty-seven grammar school children, paying special attention to the cases of very high intelligence. Campbell (1952) stressed the influence of home environment on the educational progress of selective secondary school children, and the Central Advisory Council for Education (1954) emphasized the role of the home in preventing children from making the most of their opportunities. There is a need for controlled studies in this field, and it is appreciated that clinical studies must be supplemented by rigorously designed enquiries. The analysis of case-studies can, however, be the source of valuable insights (Argyle, 1957), though it is often difficult to systematize these from clinical records designed for practical purposes, and the more subtle emotional factors do not easily lend themselves to statistical evaluation, as Kanner (1948) points out in discussing the effects of parental attitudes.

III.—TYPES OF CASE REFERRED.

(1) *Main Reasons for Referral.*—There is often a multiplicity of complaints from the source of referral with regard to an individual child. The following classification (Table 1) attempts to present the main reasons for referral to the Child Guidance Centre.

TABLE 1

MAIN REASONS FOR REFERRAL TO CHILD GUIDANCE CENTRE.

Reason	Boys	Girls	Total
Difficult behaviour	9	7	16
Very poor work	8	3	11
Truancy or refusal to attend school	4	6	10
Stealing ; dishonesty	6	3	9
Peculiar or withdrawn behaviour	3	3	6
Anxiety symptoms	1	1	2
Difficulty with speech	1	1	2
Enuresis	1	1	2
Soiling	1	0	1
Excessive sexual interest	0	1	1
Totals	34	26	60

Of the sixteen cases referred for difficult behaviour, three were difficult only at home, and three only at school. Thirty-four of the children were reported by their school to have an unsatisfactory standard of work. Twenty-three of the total group stated that they disliked attending a grammar school.

(2) *Age of Referral.*—The average age of the group was 14 years (boys 14 years 3 months, girls 13 years 9 months).

TABLE 2

YEAR OF SCHOOL CAREER IN WHICH CHILDREN WERE REFERRED.

Year of School Career	Number of children Referred
1st	9
2nd	10
3rd	19
4th	15
5th	4
6th	3
Total	60

It will be seen from Table 2 that the peak ages of referral are in the third and fourth years of the grammar school career. This peak may be due (i) to the onset of puberty, these years being notorious for their troublesomeness to teachers (Wall, 1948) or (ii) to the children being given the opportunity to settle down before being brought forward as child guidance cases or perhaps (iii) to the school staffs considering referral to a child guidance clinic as a last resort.

IV.—MAIN CAUSATIVE FACTORS.

It is common experience of child guidance workers to find that a number of factors have combined to produce the symptoms of emotional maladjustment in any one case; and that only very rarely is there a clear-cut relationship between a particular factor and the symptoms of maladjustment. This is exemplified in the case studies quoted in the Appendix as well as by the considerable overlapping of factors in Table 3.

The case material has been analysed to show the relative importance of the following factors: (*a*) Physical Condition; (*b*) Level of Intelligence; (*c*) Personality Defect or Disturbance; (*d*) Parental Attitudes; (*e*) Sibling Rivalry; (*f*) Home Conditions and Family Circumstances. The incidence (summarised in Table 3) and effects of these factors are described below.

TABLE 3

THE INCIDENCE OF THE MAIN CAUSATIVE FACTORS.

	Factor	No. of children affected*	
(*a*)	Physical Condition:		12
	(i) Sensitivity to physical handicap or abnormality	4	
	(ii) Frequent ill-health	3	
	(iii) Delicate state of health	2	
	(iv) Visual defects of marked degree	2	
	(v) Traumatic physical experience	1	
(*b*)	Intelligence Level in Borderline Range		14
(*c*)	Personality Defect or Disturbance		15
(*d*)	Faulty Parental Attitudes		42
	(i) Constraint (strictness, over-anxiety)	16	
	(ii) Parental rejection	11	
	(iii) Lack of harmony between parents	6	
	(iv) Over-indulgence	5	
	(v) Over-casual attitude	3	
	(vi) Neglect	1	
(*e*)	Sibling Rivalry		10
(*f*)	Unsatisfactory Home Conditions and Family Circumstances		24
	(i) Broken home	15	
	(ii) Economic stress	13	
	(iii) Serious physical illness in the close family circle	8	
	(iv) Mental illness in family	13	

In only four cases was there no obvious environmental factor. In as many as fifty-one cases the emotional atmosphere of the home was unsatisfactory. In twenty-four cases there were adverse home conditions and family circumstances. In twenty cases both the psychological atmosphere and the material conditions seemed responsible for the child's maladjustment. The ways in which the emotional climate and stability of the home affect the child's learning and adjustment in school have been described in recent literature (Peel, 1956).

---

* It should be noted that several factors are usually operative in an individual case.

(*a*) *Physical Conditions.*—From Table 3 it will be seen that physical factors do not seem to play a large part in the case-histories of these children. It should be realized, however, that missed schooling, even for a short period, can seriously retard progress at this level. There is usually little opportunity for making up for time lost, and the child who is prone to frequent illnesses suffers out of all proportion to the nature and extent of his physical disability or incapacity. There is also little special provision throughout the country for physically handicapped (other than the blind and the deaf) pupils of superior intelligence, and there appears to be a need for residential and day selective schools, mostly perhaps on a regional basis, for such pupils.

(*b*) *Level of Intelligence.*—Individual assessment of intelligence level is important even in the case of these children, who have already been selected for a grammar school curriculum, as one of the questions to be answered is whether the child is finding grammar school work too much of a burden. As Wall (1948) points out, there are some children who, finding the work beyond their capacity, react either by overt behaviour disorders or by so-called laziness or day-dreaming. Burns (1949) had a questionnaire answered by head teachers of Birmingham grammar schools which indicated that the most frequent factor in failure to complete the school career satisfactorily was "lack of sufficient intelligence." A British Psychological Society Inquiry edited by P. E. Vernon (1957) stresses the complexity of the considerations involved in secondary school selection procedures and points out that there may be decisions on borderline cases which lead to frustration for the individuals concerned. In areas which have a large grammar school entry, it may well be the case that more children are admitted to such schools than can cope with the curriculum without strain—a possibility which is not sufficiently recognized.

Fifty-two of the children in this study were individually tested on the Terman-Merrill Scale (Form L). Some were given additional tests, but Table 4 gives the intelligence level of those tested in terms of the Terman-Merrill I.Q.

TABLE 4

INTELLIGENCE LEVEL OF THE FIFTY-TWO CHILDREN TESTED ON THE TERMAN-MERRILL SCALE (FORM L).

I.Q. range*	Boys	Girls
110—119	6	8
120—129	6	10
130—139	11	2
140—149	4	2
150 +	3	0
Total	30	22
Mean I.Q.	131·3	123·1
σ	12·65	8·8

* The McNemar adjustment has been made where applicable.

The mean I.Q. of the whole group was 127·2, with a S.D. of 11·95. Two girls were tested on the Wechsler-Bellevue Scale (I.Qs. of 120 and 136). One girl was unco-operative with testing.

Children with an I.Q. of at least 120 are generally considered as definitely suitable for a grammar school course, while some with lower I.Qs. achieve success. It will be seen from Table 4 that only about a quarter of the children (those in the I.Q. 110–119 range) had even the suggestion of being intellectually unfit for a grammar school curriculum. Of the eleven children specifically referred because of ' very poor work,' only three had I.Qs. below 120, and five had I.Qs. of 130+.

Five children, although of adequate or high intelligence, were having difficulties because of comparatively low attainments in one of the basic subjects. For example, Case 5 (I.Q. 120) had trouble with mathematics; Case 49 (I.Q. 129) was " reasonably good in mathematics and science, but could not put things down on paper " ; and Case 57 (I.Q. 147) was good in oral work, but had great difficulty in expressing himself in writing. In these cases remedial help was given at the Centre, and it is worth emphasizing that the remedial approach need not be confined to the pupils of non-selective schools.

A number of children seemed to be unable to face the increased intellectual competition which they found at the grammar school. Case 51 (I.Q. 127) had been considered as the star of his primary school, but was an ordinary member of the grammar school. Case 12 (I.Q. 120) was a physically handicapped girl whose intellectual superiority in the primary school had compensated for a congenital deformity, but who broke down in the grammar school.

(*c*) *Personality Defect or Disturbance.*—There is invariably considerable interaction between the personality of the child and his environment. It is not easy, therefore, to decide whether defects or disturbances of personality are cause or effect in individual cases ; but a distinction may be attempted between those cases where the problem appears to result from the reaction of an intrinsically normal child to a disturbing situation and those cases where there is evidence of marked personality defect or disturbance. In the former case, the child is temporarily disturbed and will quickly make a better adjustment in improved circumstances ; in the latter, the disturbance is usually long-standing, deep-seated and not easily treated.

In this group of sixty cases, no more than fifteen could be placed, on the basis of the psychiatrist's assessment, in the second category. Of these, fifteen, eleven were described as having ' schizoid tendencies,' and four as being ' possibly psychopathic.' No case was considered in need of hospital treatment.

Some children showed evidence of poor character development in that they were not prepared to sacrifice the pleasures of the present for long-term aims. Case 58 was a boy with considerable ability but no sense of values. He was entirely pleasure-seeking, and lived for the moment, without regard for his future career. Such an attitude among some older pupils presents a real problem, aggravated by the fact that they may have friends who are already at work and consider themselves adults.

(*d*) *Parental Attitudes.*—The role of parental attitudes in the causation of maladjustment among children has been stressed by many writers (e.g., Kanner (1948), Lewis (1954) and Stott (1956)). In forty-two cases, faulty or unsatisfactory parental attitudes militated against the satisfactory development and progress of the child. The nature of these attitudes, as assessed by the clinic team, is shown in Table 3.

The parents of those children who presented nervous or neurotic symptoms (twenty-two cases) tended to be over-anxious, there also being much parental conflict over handling the child. The most significant attitude of the parents of the children who were delinquent in some way, or who played truant (nineteen cases in all) was a lack of interest, supervision, and control. These children were rarely given much encouragement or support at home for their efforts in school. Where the children were aggressive and troublesome, but non-delinquent (sixteen cases), the atmosphere of the home was often repressive. Not infrequently, one parent was over-strict and authoritarian, while the other was over-protective and shielding. Authoritarian attitudes often lead to resentment and rebellion, especially during puberty. Pressure and nagging tend to lead to a lack of effort on the part of the child.

In one family, there was a conflict of strong personalities over religious matters, which created an atmosphere of tension. Case 13 was not allowed by her father to aim at becoming a nurse—her one goal in life.

(*e*) *Sibling Rivalry.*—In ten cases there was definite jealousy of one or more siblings, which often led to attention-seeking and demanding behaviour. In three of these cases a sibling was physically handicapped and required a great deal of the parents' attention ; one child was unfavourably compared by his parents with one of his siblings ; and Case 23 was always protected and excused by his twin sister.

(*f*) *Home Conditions and Family Circumstances.* (i) *The Broken Home.*—The lack of a cohesive, complete family unit is often cited as responsible for maladjustment and delinquency (Lewis, 1954 ; Cameron, 1947). In this group, the home was broken or incomplete in fifteen cases, in which the deficit in the family unit was as follows:

Father dead	7
Mother and father separated	3
Mother dead	2
Both parents dead	1
Father away from home for long periods	1
Child illegitimate ; father not member of family	1

In the cases where the parents had separated, there had been a background of domestic stress and conflict beginning in the child's infancy. The main effects of the broken home were :

(1) The child was forced to spend a period in an institution (two cases), or to suffer many changes of residence (two cases).

(2) The loss of a parent through death or separation meant a heavy burden on the remaining parent and in some instances a part of the burden fell on the child (six cases).

(3) Some children were looked after by relatives (four cases), sometimes reluctantly.

Apart from the direct psychological effects of the broken home—feelings of insecurity, strain, and unsettledness—the material hardships which often result should not be underestimated.

(ii) *Economic Stress.*—In thirteen families there was definite economic stress. Of these, in seven cases there was financial strain in maintaining the standard of the home ; and in six cases the social worker reported a standard of living amounting to poverty. In addition, two families, although not poor, lived in a very poor area, from which few grammar school children came.

There were three main results of such circumstances :

(1) In at least ten cases, the child had marked feelings of social inferiority, with consequent dislike, and even avoidance in some instances, of going to a grammar school. These feelings were related mainly to a lack of sufficient

pocket-money, and sensitivity concerning the material standard of the home and inferior clothes. These children felt ' out of it ' both at school and in their own neighbourhood, where the local children might regard them as ' snobs'.

(2) In eight homes there was over-crowding and a consequent lack of facilities for study. Some of the children had to attempt their homework in the presence of noisy siblings, others in rooms lacking warmth. Case 10 (a boy of I.Q. 147) did his homework in the front room in the summer, but in the winter this room had no fire, and he had to wait to work in the living room, which was not available for him until late in the evening.

(3) In all thirteen cases, there was a general atmosphere at home of strain and pressure, sometimes with a premature burden of responsibility for the child. In at least two cases (both boys), there were many demands on their time to help with the baby and the housework. Case 18 had been an unwanted child at first because the father was unemployed.

(iii) *Serious Physical or Mental Illness in the Close Family Circle.*—It is rarely easy for complete and reliable information to be obtained on the physical and mental health of the family circle. It was found that in at least thirteen cases one member of the family suffered from mental illness, and in eight cases from physical ill-health.

The main effects of these circumstances were :

(1) The atmosphere of the home might become morbid, gloomy, or anxiety-ridden.

(2) The absence of a parent from the home, due to long periods in hospital, adversely affected the domestic routine.

(3) The child might be prevented from giving adequate attention to homework as a result of the unstable behaviour of a member of the household.

V.—METHODS OF TREATMENT.

Treatment is considered here in its widest sense. A child guidance team thinks in terms of trying to improve the situation at home or at school, and of attempting to modify attitudes *towards* the child as well as of direct treatment. The gravity of the problem is occasionally lessened by the mere referral of the case to a clinic and the discussions and investigations which ensue.

The level of intelligence of this group of children, and usually also of their parents, makes advice and individual psychotherapy of a verbal nature more effective and applicable than such treatment often is in a number of cases routinely referred. In all cases, advice was given to the parents and an attempt was made to give them insight into the nature of the problem when this was lacking. In thirty-seven cases individual psychotherapy with the child was the main approach, in some cases in addition to other means (such as drugs, necessary in two cases only) ; in five cases remedial help was given, as previously mentioned ; one child joined a play therapy group ; in two cases, treatment was confined to social work with the family. In fifteen cases, general advice to the source of reference and appropriate recommendations with regard to schooling were sufficient.

In order to avoid undue interruption of their school work and thus exacerbating the problem in some cases, every effort was made to see grammar school children at the Centre during school holidays, unless more intensive treatment was necessary.

Personal discussions between the educational psychologist and the Head (and often the teachers concerned) took place whenever possible. Forty-four of the children were able to continue at their grammar school. In the other sixteen cases a change of school was made (thirteen on the recommendation of the Centre), as follows:

Transfer to secondary modern school	5
Change to another grammar school	2
Placement in residential school	7
Transfer to private school	2

The appropriate residential placement of maladjusted children of high intelligence is very difficult; there are still few schools which cater for their needs. Only rarely do residential schools for maladjusted children offer a curriculum at the grammar school level, and, as the Underwood Report (1955) points out, it is most important that these children should have an environment where their varied needs are understood and met.

Hospital treatment was not considered necessary in any of the sixty cases.

VI.—OUTCOME OF TREATMENT.

Twenty of the children were satisfactorily adjusted while remaining at a grammar school, and thirteen considerably improved. The five children who were transferred to a secondary modern school all adjusted well there, without the need for any direct treatment. In five cases little or no progress was made, three of these failing to co-operate in the suggested treatment. Nine children left school not long after being seen at the Centre, whereupon they no longer came under the aegis of the Local Education Authority. One child was still attending the Centre at the time of the completion of this study.

Of the seven children who went to residential schools, three were subsequently well adjusted, one boy going on to a university after returning to a day grammar school. One case improved; two children were dealt with by the Juvenile Court, and committed to Approved Schools; and in the other case, the outcome is not known.

Information about the post-school careers of some of these children would be very valuable, but this is rarely available.

ACKNOWLEDGEMENTS.—The writer would like to thank Dr. G. Stuart Robertson, Deputy Principal School Medical Officer, Liverpool, as well as Dr. H. S. Bryan and Dr. I. Leveson, Consultant Psychiatrists, Liverpool Education Committee Child Guidance Centre, for their kind permission to publish the data of this study.

VII.—REFERENCES.

[1] ARGYLE, MICHAEL (1957). *The Scientific Study of Social Behaviour*. London: Methuen.

[2] BURNS, C. L. C. (1949). Maladjusted children of high intelligence. *Brit, J. Educ. Psychol.*, 19, 137-141.

[3] CAMERON, NORMAN, A. (1947). *The Psychology of the Behaviour Disorders*. Boston: Houghton Mifflin Company.

[4] CAMPBELL, W. J. (1952). The influence of home environment on the educational progress of selective secondary school children. *Brit. J. Educ. Psychol.*, 22, 89-100.

[5] CENTRAL ADVISORY COUNCIL FOR EDUCATION (1954). Report on "*Early Leaving*."

[6] H.M.S.O. (1955). *Report of Committee on Maladjusted Children*.

[7] KANNER, LEO (1948). *Child Psychiatry*. Oxford: Blackwell.

[8] LEWIS, HILDA (1954). *Deprived Children*. Oxford University Press.

[9] MANNHEIM, H. (1948). *Juvenile Delinquency in an English Middletown*. Cambridge University Press.

[10] MCCLELLAND, WILLIAM (1942). *Selection for Secondary Education*. University of London Press.

[11] PEEL. E. A. (1956). *The Psychological Basis of Education*. Edinburgh: Oliver and Boyd.

[12] SCHONELL, F. J. (1952). The development of educational research in Great Britain. Part VII—Maladjusted Pupils. *Brit. J. Educ. Psychol.*, 22, 30-44.

[13] STOTT, D. H. (1956). *Unsettled Children and their Families*. University of London Press.

[14] TERMAN, L. M. (1926). *Genetic Studies of Genius*. Stanford University Press.

[15] WALL, W. D. (1948). *The Adolescent Child*. London: Methuen.

[16] VERNON, P. E. (ed. 1957). *Secondary School Selection: A B.P.S. Inquiry*. London: Methuen.

VII.—APPENDIX.

CASE STUDIES.

Case 1 (a girl of 12) was referred to the Centre during her first year at a grammar school. She began to shrink from contact with the outside world, and could not face the thought of going out by bus, being in shops, or of going to school. She had feelings of remoteness, and was very moody. There seemed to be many factors operating in this case. Intellectually, she was a border-line pupil (Terman-Merrill I.Q. 110). Her symptoms were exaggerated following the death of her father, to whom she had been closely attached. Both parents had a history of emotional disturbance. There was poverty and overcrowding at home, where it was difficult to find quiet and room to do homework.

After a prolonged period away from school, and psychiatric treatment, the girl went to a secondary modern school. She did not wish to go back to a grammar school, feeling that she was now far behind with her work and being anxious to return to her friends at the secondary department of the school she had attended before transfer to the grammar school. She settled down well at the secondary modern school.

. . . . . .

Case 6 (a boy of 13, with a Terman-Merrill I.Q. of 147) was referred during his second year at grammar school. His headmaster did not consider that he was justifying his place at a grammar school, either from the point of view of work or behaviour. The boy himself was not happy at school because he had made no friends there and because the other boys always had far more pocket money than he had. He stole money from home, and ran away on several occasions.

Examination showed him to be of very superior intelligence, but he had a marked sense of social inferiority in relation to the poverty of his home. His father had died seven years previously, and his mother had to work long hours with the result that there was inadequate supervision of the boy. The mother herself was worn, depressed and anxious. She had some responsibility for her own aged and ailing parents. The boy was made to stay at home to keep his mother company and to help in the house, and he resented his lack of opportunity to take part in outside interests. There had been an early traumatic experience.

He went to a residential school in the country, but while there got into trouble with the police and was committed to an Approved School.

. . . . . .

Case 48 (a boy of 14, with a Terman-Merrill I.Q. of 155) was referred for examination because of his persistently difficult behaviour in school. He was sullen and morose, defiant and disobedient, although said to be well-behaved at home. The material conditions of the home were good, but there was much conflict between the parents over handling the boy. The mother was indulgent to him, although she herself had a generally aggrieved attitude ; the father was very authoritarian in manner. The boy suffered from frequent attacks of bronchitis, and had missed some schooling as a result. In addition, he had been quite outstanding at the primary school stage, and since his transfer had resented the competition he found in the grammar school.

The mere fact of his referral for examination to a child guidance centre had an effect on his behaviour, and he responded quickly to help and advice. He was able to complete his grammar school course with success.

THE VALUE OF TEACHERS' RATINGS OF THE ADJUSTMENT OF THEIR PUPILS*

BY J. G. THORPE

(*Banstead Hospital, Sutton, Surrey.*)

SUMMARY. The present experiment was designed to test the relationship between teachers' ratings of adjustment and maladjustment, and more objective estimates obtained by psychological testing. Three groups of thirty-six children were built up—a group of most adjusted, a group of least adjusted, and a group selected at random.

Objective tests of neuroticism which had been shown previously to distinguish between neurotic and normal children were found, with the exception of the Sociometric test, not to differentiate between these three groups.

The results indicated that, if the battery of neuroticism tests is an acceptable criterion of neuroticism, the teachers' ratings of adjustment bore no relation to this criterion. It may be argued, however, that the criterion was an inadequate indicator of psychiatric abnormalities and, therefore, that it did not follow that teachers were unable to detect these abnormalities.

The teachers in this study appeared to be leaning heavily upon the Sociometric status of their pupils for their judgments, a knowledge of which they were bound to have from everyday observation. It is argued that the Sociometric test discriminated between the groups of children selected by the teachers probably by virtue of its projection on to some dimension other than neuroticism. This dimension is possibly that of introversion-extraversion.

I.—INTRODUCTION AND PROBLEM.

AN integral part of a child's adjustment to his environment is his adjustment at school, if only because of the number of hours he spends there each day. In many cases, the parents have little or no idea of how their children are faring at school, except in terms of their position in class at the end of each school year, and even this is complicated by the almost standard procedure of streaming. Assessments of the child's emotional and social adjustment, however, rarely find their way into the school report so that the parents are given no indications as to how their children are developing in this respect. The failure even to make reference to these aspects of adjustment no doubt results, in some cases at least, in the view that school achievement is the *raison ȇ'etre* of the child's education.

Those parents who regard education as including emotional and social development in addition to academic achievement, must be satisfied, in the absence of school assessments, with their own assessments of their children, but many must realise that they have no idea about how other children behave in this respect. They have no norms with which to compare their own children. This deficiency would to some extent be remedied if assessments could be made at the school by teachers, whose albeit subjective norms would be based on personal experience of large numbers of children in an almost standardised testing situation—the school classroom, *provided that these assessments can be shown to have sufficiently high validity*.

Although it is true that a number of children find their way to child guidance clinics as a result of difficulties to which the school authorities have drawn attention, for example truanting or unmanageable school behaviour, it is the thesis of this paper that if children behave themselves at school, and give the

---

* This work was carried out at the Maudsley Hospital, under the direction of Professor H. J. Eysenck, and financed by The Central Research Fund of the Bethlem Royal and Maudsley Hospitals.

impression of doing their best in their lessons, their problems of adjustment will go unnoticed, and, therefore, untreated. In other words, we are setting up the hypothesis that teachers' ratings of adjustment will be shown to have little value when compared with objective estimates obtained by psychological testing.

Work published elsewhere (Thorpe and James, 1957; Thorpe, 1958) suggests that the Maudsley battery of neuroticism tests distinguishes between normal and neurotic children in a similar way to that in which normal and neurotic adults are differentiated. In the first of these two papers the results indicated that the influence of the testing situation was more marked than had hitherto been supposed. Nevertheless, the performance tests of neuroticism tended to discriminate between the two groups even when the testing situation was carefully controlled by testing all the normal and neurotic children in their schools.

In the second of these papers the analysis was continued in an attempt to isolate a factor of neuroticism from these new data uncontaminated by the testing situation. Although a centroid analysis followed by rotation to simple structure was unsuccessful in this respect, a more positive conclusion was suggested as a result of the application of Criterion analysis (Eysenck, 1950). The factor finally isolated by this method showed a correlation of ·492 between the discriminatory powers of the tests and the saturations of the tests on the factor. It was argued that this correlation was statistically acceptable and this conclusion was supported to some extent when the factor scores, which were calculated roughly, showed discrimination between the two groups at the 20 per cent. level of significance.

The conclusions drawn from these two papers were that in spite of a rigorously controlled testing situation, the objective performance tests of neuroticism still showed signs of discriminating between the normal and neurotic groups in the expected manner, although this conclusion did not apply to the verbal (questionnaire) tests. In addition a factor of neuroticism could be isolated from the data which in many ways resembled the factors isolated in studies of adults.

In the present paper the same battery of neuroticism tests employed in the two studies above were used as a criterion against which the teachers' ratings of adjustment and maladjustment were to be evaluated.

II.—METHOD.

(*a*) *Pupils.*—Three L.C.C. schools, a county secondary school for boys, a secondary school for girls, and a mixed secondary modern school, agreed to co-operate in this project. Within each school three classes were chosen with the head's assistance. The only requirement was that no class must comprise the problem and backward children of the school, as previous experience had indicated that such a class would in all probability be too homogeneous in respect of maladjustment.

The Form teachers were then asked to select from their class the four most and the four least adjusted pupils. No advice regarding criteria of adjustment were given, but each form teacher was asked to write a brief note giving his reasons for allocating each child to the category chosen. Four more children were chosen from each class on a random selection basis. Each class, therefore, provided four pupils rated as ' most adjusted', four as ' least adjusted', and four selected at random. The total sample comprised thirty-six children in each of the three groups.

(*b*) *Tests.*—All children were given Thurstone's PMA test of intelligence (Thurstone and Thurstone, 1947), together with questionnaire tests described in detail in the two books by Eysenck (1947, 1952), namely :

(1) Sentence completion.
(2) Annoyances.
(3) Ways to be different.
(4) Medical questionnaire.
(5) Worries and anxieties.
(6) Interests.
(7) Sociometric (described in detail by Thorpe, 1955).

The above tests were given in group form to all pupils in the twelve classes.

The thirty-six selected children were then tested individually on the following tests. Testing was carried out in the mobile laboratory described in an earlier paper (Thorpe and James, 1957) and adequate descriptions of the tests can also be found in Eysenck's books. These were :

(8) Manual dexterity.
(9) Track tracer.
(10) Leg persistence.
(11) Dynamometer persistence.
(12) Static ataxia.
(13) Body sway suggestibility.
(14) Body sway reversals.
(15) Autokinetic movement.
(16) Rail walk.
(17) Tapping (max.).
(18) Reaction time. Regular preparatory interval.
(19) Reaction time. Irregular preparatory interval.
(20) Luria. Left-hand disturbance.
(21) Dark adaptation.
(22) Mirror drawing.

III.—RESULTS.

The test means and standard deviations obtained by the three groups are given in Table I, together with the " F " ratios and significance levels.

TABLE 1

Test	Expected direction of score indicating neuroticism	Most Adjusted		Random		Least Adjusted		" F "	Probability	Differences in Means between most and least adjusted in expected direction
		Mean	*Variance*	*Mean*	*Variance*	*Mean*	*Variance*			
Age	—	159·867	147·357	158·916	150·939	157·815	160·397	·519	N S*	—
Sex	—	19 male 17 female		18 male 18 female		20 male 16 female				
PMA	—	137·342	900·123	135·000	981·502	130·316	1019·227	·413	NS	—
1. SC	High	4·842	5·704	5·222	4·806	4·974	3·972	·189	NS	Yes
2. A	High	23·605	58·191	22·750	62·479	22·053	41·781	·492	NS	No
3. WTBD	High	14·553	102·092	14·083	53·564	15·184	70·317	·133	NS	Yes
4. MQ	High	8·526	19·878	6·389	8·587	7·474	12·797	1·529	NS	No
5. W and A	High	11·263	24·415	10·278	25·235	11·763	25·267	·421	NS	Yes
6. I	High	15·184	37·344	12·306	24·390	13·632	24·077	1·521	NS	No
7. S	Low	·378	·044	·207	·032	·132	·058	13·460	·1%	Yes
8. MD	Low	210·130	374·975	204·165	347·850	200·660	459·700	1·401	NS	Yes
9. TT	High	86·710	2055·775	101·250	3880·525	104·735	2937·775	1·176	NS	Yes
10. LP	Low	44·133	2351·142	41·751	890·937	46·185	1753·560	·459	NS	No
11. DP	Low	62·208	3464·604	82·836	12811·500	86·370	10713·492	1·235	NS	No
12. SA	High	·756	·650	·507	·348	·487	·328	1·260	NS	No
13. BSS	High	2·382	9·763	2·035	6·811	2·296	4·466	·716	NS	No
14. BSR	High	8·579	89·548	14·444	110·083	10·737	156·902	1·566	NS	Yes
15. AM	High	13·737	83·388	11·556	42·825	11·921	53·318	·795	NS	No
16. RW	Low	93·158	367·488	88·888	533·132	85·316	648·708	·681	NS	Yes
17. TM	Low	60·000	131·081	55·306	61·018	57·053	90·376	1·635	NS	Yes
18. RT reg.	High	28·763	79·267	28·444	54·311	29·053	42·051	·058	NS	Yes
19. RT irreg.	High	31·684	104·060	31·472	39·913	31·816	40·533	·335	NS	Yes
20. L	High	2·974	18·026	1·861	6·523	2·316	5·735	1·214	NS	No
21. DA	High	18·474	10·743	18·306	10·504	18·053	14·267	·740	NS	No
22. MD	High	40·736	4419·328	68·888	9429·136	27·368	1547·696	3·270	5%	No

* NS.—Not significant.

IV.—DISCUSSION.

To summarise Table 1, out of twenty-two measures of neuroticism, only two show statistically significant differences between the three groups of children in this study. These two exceptions are the Sociometric Test and the test of mirror drawing. In the case of the latter, however, as the direction of the means is different from expectation, this statistically significant result is probably spurious. We can conclude, therefore, that only the Sociometric test succeeds in distinguishing between groups of most adjusted, least adjusted, and randomly selected pupils. It will also be seen from this Table that, out of the seven questionnaire tests, four of them show differences in the expected direction between the most and least adjusted groups, while out of fifteen objective performance tests, seven of them behave in this manner. These findings are almost exactly what would have been predicted on the basis of chance alone. It seems clear, therefore, that with the exception of the Sociometric test, the results for which are clearly and significantly in line with prediction, the remainder of the tests do not give results corresponding with the judgments made by teachers.

If we accept our test battery as a satisfactory criterion of neuroticism, therefore, we must conclude that teachers' ratings of adjustment bear no relation to assessments of neuroticism obtained in this manner. It could be argued, however, that as the correlation of the test battery with the normal-neurotic criterion in the original studies (*loc. cit.*) was less than ·5, it would still be possible for the teachers' ratings of maladjustment to correlate significantly with the normal-neurotic criterion if the experiment could be carried out.

It is interesting to note from Table 1 that the teachers as a group were successful in not allowing age and intelligence to influence their judgments—the three groups are very similar indeed in respect of these two variables. It would appear, however, that the Sociometric status or popularity of the pupils could have had some influence upon the teachers' judgments, a conclusion which is borne out when we consider the teachers' criteria of adjustment. Five of the twelve class teachers, selected on a random basis, produced the following lists which are given below :

Class No. 1.

Most Adjusted Children characterised by :
 Happiness, good powers of leadership, alertness, good schoolwork, and willingness to work.
Least Adjusted Children characterised by :
 Lack of co-operation, bad behaviour, anxious to be noticed, sulkiness, shyness, and reserve, and quietness.

Class No. 2.

Most Adjusted were :
 Interested in work and play, had co-operative parents, were good leaders.
Least Adjusted were :
 Nervy and erratic, poor at schoolwork. One child suffered from nervous deafness, another stole occasionally.

Class No. 3.

Most Adjusted were :
 Capable, self-reliant, good mixers, good workers, interested in work, and not easily ruffled.
Least Adjusted characterised by :
 Untidy work, fidgeting, shyness, lack of helpfulness.

Class No. 4.

Most Adjusted were :

Quiet, keen and effervescent, helpful, and reliable.

Least Adjusted were :

Resentful of correction, lazy, sulky, lacking in energy, and backward.

Class No. 5.

Most Adjusted characterised by :

Contentedness, popularity, friendliness.

Least Adjusted were :

Leaders, truculent, difficult to get to know, fidgety, bad tempered, spineless, lethargic, and babyish.

It is fairly clear from these that Sociometric status is probably one of the factors upon which the teachers' ideas of adjustment are based.

Other findiugs also require explanation. In our previous papers (*loc. cit.*) we showed that the Sociometric test discriminated between children attending a child guidance clinic and normal children at a satisfactory level of statistical confidence, as well as did some of the other tests of neuroticism. These findings taken in conjunction with the present ones indicate that the Sociometric test distinguishes between neurotics and normals, and also between groups of most and least adjusted pupils selected by school teachers. The tests of neuroticism, on the other hand, distinguish only between the neurotic and normal children. A possible explanation of this would appear to be that the Sociometric test discriminates between neurotic and normal children by virtue of its projection on to the neuroticism axis, while it discriminates between most and least adjusted pupils on the basis of its projection on to some other dimension, possibly that of introversion-extraversion. This would mean that the pupils high in Sociometric status occupied positions at the normal end of the neuroticism continuum and/or were extremely extraverted. Those low in Sociometric status would fall at the neurotic end of the neuroticism continuum and/or would be extremely introverted. Thus, two groups of children equated on introversion could show significant differences in Sociometric status because of a difference in neuroticism, while groups equated on neuroticism could show Sociometric differences because of differences in introversion-extraversion. This explanation would appear to be a reasonable one, though it should be regarded at this stage as an hypothesis for further investigation.

ACKNOWLEDGEMENTS.—Thanks are due to the L.C.C. and the heads and teachers of the schools concerned for their valuable co-operation, as well as to the Central Research Fund of the Bethlem Royal and Maudsley Hospitals for financial support.

V.—REFERENCES.

[1] EYSENCK, H. J. (1952). *The Scientific Study of Personality*. Routledge and Kegan Paul.

[2] EYSENCK, H. J. (1947). *Dimensions of Personality*. Kegan Paul.

[3] EYSENCK, H. J. (1950). Criterion analysis—An applicatisn of the hypothetico-deductive method to factor analysis. *Psychol. Rev.*, 57, 38-53.

[4] THORPE, J. G. (1955). An investigatisn into some correlates of sociometric status within school classes. *Sociometry*, 18, 49-61.

[5] THORPE, J. G. (1958). Neuroticism in children : II—Identification and measurement of a neuroticism factor. To appear.

[6] THORPE, J. G., and JAMES, D. P. (1957). Neuroticism in children : I—An investigation of normal and neurotic group differences. *Brit. J. Psychol.*, 48, 26-34.

[7] THURSTONE, I. L., and THURSTONE, T. G. (1947). *S.R.A. Primary Mental Abilities*. Chicago : Science Research Associates.

TWO PERSONALITY CHARACTERISTICS RELATED TO ACADEMIC ACHIEVEMENT

By R. LYNN

(*Exeter University.*)

Summary. Experimental studies suggest that capacity for sustained work depends largely on an individual's level of drive and rate of accumulation of inhibition. This work implies that good academic achievers should be characterized by high drive levels and a slow rate of accumulating inhibition. The Maudsley Personality Inventory measures these two personality dimensions and a comparison of the scores of university students with those of controls confirms both predictions.

I.—Introduction.

The theories of behaviour built up by experimentalists (e.g., Hull, 1952) have now reached a stage where there is broad agreement on certain general principles. At this point the findings of laboratory experimentation should be of considerable use to the various fields of applied psychology, and, in fact, considerable use of behaviour theory as an explanatory device has been made in the areas of personality development (e.g., Dollard and Miller, 1950 ; Sears, Maccoby and Levin, 1957), psychotherapy (Davis, 1957 ; Eysenck, 1957), and individual differences (Eysenck, 1957). As yet, however, and with one notable exception (Peel, 1956), rather little use of behaviour theory has been made in educational psychology. The present paper is concerned with two predictions from behaviour theory to the problem of individual personality differences in capacity for academic work.

It is clear that people differ considerably in their capacity for sustained and concentrated work and it seems likely that this personality characteristic contributes to good educational attainment, perhaps to a considerable degree. In behaviour theory terms, sustained work is largely dependent on two factors, namely, the strength of drive and the accumulation of inhibition as work proceeds. Individuals with a capacity for sustained work should, therefore, be characterised in the following way : (1) they should have high drives ; (2) they should accumulate reactive inhibition slowly with continuous work. Some of the recent studies of Eysenck (e.g., 1957) simplify the testing of these predictions. In this work, Eysenck has identified his personality dimension of neuroticism with autonomic drive, and that of extraversion with the fast accumulation and slow dissipation of reactive inhibition. There is now a considerable amount of experimental evidence supporting this theory, for which the reader is referred to the original reports (e.g., Eysenck, 1957). If Eysenck's theory is accepted, the predictions about educational attainment can be stated as follows : good educational attainers should (1) score high on neuroticism ; (2) low on extraversion.

Several recent studies have produced evidence supporting this theory. The most extensive of these is that of Furneaux (1957), who has shown that students who do well at university score more highly on neuroticism and lower on extraversion. He also puts forward the interesting view that extraversion only begins to have a detrimental effect on educational attainment at the university level. His argument on this aspect of the question is as follows : if candidates who are accepted and rejected for university places are considered, their extraversion scores are very similar. However, since their educational attainment differs (this being largely the basis of acceptance or rejection), it appears that at the level of university entrance extraversion is not related to educational attainment. Furneaux suggests that the explanation for this lies in the stricter supervision of school life in which the tendency of the extravert to dissipate his energies is held in check.

A corrobatory study of the introversion-extraversion finding has been reported by Broadbent (1958). Students graduating at Cambridge were divided into those obtaining good and poor degrees, and their level of extraversion assessed by means of the triple tester; this test showed that students who do well were significantly more introverted than those who do badly. This study also showed that the two groups of students did not differ in intelligence as assessed by the A.H.4 test and suggests, therefore, that introversion-extraversion acts independently of intelligence in affecting educational attainment.

Several studies have been made of the relation of educational attainment to anxiety, with rather conflicting results. The concept of anxiety is rather an unsatisfactory one, since although it is largely a measure of neuroticism (autonomic drive), it is also related to introversion. Hence, the theory would predict that high anxious subjects should do well in tasks where sustained work is required, although it is not certain how far this is due to neuroticism or introversion. Some evidence in support of this prediction has been presented elsewhere (Lynn, 1955).

The present paper reports findings extending the work of Furneaux and Broadbent. Essentially, these investigators have shown that good students differ from poor students in the expected directions of introversion and neuroticism. The theory should also predict that students as a whole should differ from other young people on these two dimensions.

II.—The Investigation.

Levels of neuroticism and extraversion were assessed in university students and controls by means of the Maudsley Personality Inventory (see Eysenck, 1956). University students were all in their first year at university; mean age of women$=18\cdot8$, mean age of men$=19\cdot2$. Controls used were (*a*) sixty-seven female occupational therapy students of the same age (mean age$=18\cdot5$) and social background as the female university students but differing in academic motivation; (*b*) 100 male apprentices aged 16-19 years whose scores on neuroticism and extraversion were taken from a study by Field (1959). There

TABLE 1

	Norms	Apprentices	University Students
1. Men:			
Number	200	100	115
Neuroticism ..	17·8	21·2	25·5
Extraversion ..	24·6	29·3	22·4
	Norms	O.Ts.	Students
1. Women:			
Number	200	67	96
Neuroticism ..	19·4	23·8	28·2
Extraversion ..	25·2	28·5	22·9

may have been some group difference in intelligence, but since there is virtually no correlation between the two personality dimensions and intelligence, this difference is probably of little significance. However, the possibility of intelligence affecting extraversion and neuroticism scores deserves further scrutiny. The norms of the questionnaire are also used as a control, although these are probably less satisfactory because there may be some tendency for different age groups to score differently.

The results are presented in Table 1. All the differences are in the predicted directions and are significant at the ·05 level, as tested by the calculation of the standard errors of the differences between the means..

III.—DISCUSSION.

The results support the two predictions at a significant level and extend the findings of Furneaux and Broadbent. Moreover, they show that extraversion has wider detrimental effects on educational attainment than Furneaux concludes on the basis of his work and that these effects manifest themselves below the level of university extrance. This conclusion follows from the quite large differences in extraversion between the university students and the occupational therapists and apprentices. Since university entrance is obtained largely on performance in 'A' level, the results suggest that educational attainment at school is substantially affected by the introversion-extraversion dimension. Furneaux's findings of the small difference between accepted and rejected candidates is probably due to the fact that the rejected candidates must have been quite good attainers to be candidates for university entrance at all; further, the small difference he did obtain was in the expected direction, i.e., rejected candidates were more extraverted.

The finding that university students score significantly more highly on neuroticism than normal groups is perhaps surprising in view of Terman's (1925) widely accepted finding that highly talented young Americans are better adjusted than the normal population. There are now a large number of studies on this subject and the results are conflicting, so that it is becoming increasingly evident that the findings obtained depend on the measuring instrument used for assessing 'neuroticism' or 'maladjustment.' The present finding adds weight to the studies already existing which suggest that Terman's findings cannot be generalised too widely.

The findings suggest a further prediction concerning sex differences in personality and attainment. The present results confirm the common finding that women score more highly on tests of neuroticism and anxiety (e.g., Terman and Tyler, 1954). Hence, it seems likely that if intelligence and extraversion are held constant, women should, by virtue of their higher level of neuroticism, be better academic attainers than men. As far as university students are concerned, this is perhaps not an easy prediction to test. For example, the fact that at Oxford women obtain a higher proportion of first class degrees than men is doubtless due in part to greater selectivity of intake. However, a large number of studies have shown that at the age of eleven girls are significantly better attainers than boys (e.g., Yates and Pidgeon, 1957). This could be explained in terms of their higher drive level.

✓

It may be thought that the positive association of academic attainment with neuroticism presents a curious contrast with the results of laboratory experiments, in which it is generally found that high drive levels impair the learning of complex tasks. It seems that neuroticism has two different effects on attainment, a disorganising one on learning and performance in stress situations but a facilitating one in so far as it motivates sustained work. Further, at the educational levels of the university and school sixth form, its disorganising effects seem to be more than compensated for by its motivating powers. How far this is true at earlier educational levels seems to be a question well worth investigation.

IV.—REFERENCES.

BROADBENT, D. E. (1958). *Perception and Communication*. London : Pergamon Press.

DAVIS, D. RUSSELL (1957). *Introduction to Psychopathology*. Oxford University Press.

DOLLARD, J., AND MILLER, N. E. (1950). *Personality and Psychotherapy*. New York : McGraw-Hill.

EYSENCK, H. J. (1956). The questionnaire measurement of neuroticism and extraversion. *Rivista di Psicologica*, 50, 113-140.

EYSENCK, H. J. (1957). *The Dynamics of Anxiety and Hysteria*. London : Routledge and Kegan Paul.

FIELD, J. G. (1959). The personalities of criminals. Paper read to Annual Conference of B.P.S.

FURNEAUX, W. D. (1957). Report to Imperial College of Science and Technology.

HULL, C. (1952). *A Behaviour System*. New Haven : Yale University Press.

LYNN, R. (1955). Personality factors in reading achievement. *Proc. Roy. Soc. Med.*, 48, 996-998.

PEEL, E. A. (1956). *The Psychological Basis of Education*. Edinburgh : Oliver and Boyd.

SEARS, R. R., MACCOBY, E. E., and LEVIN, H. (1957). *Patterns of Child Rearing*. New York : Row Peterson and Co.

TERMAN, L. M. (1925). *Genetic Studies of Genius*. Stanford University Press.

TERMAN, L. M., and TYLER, L. E. (1954). Psychological sex differences. In *Manual of Child Psychology*. (Ed. Carmichael, L.). New York : Wiley.

YATES, A., and PIDGEON, D. A. (1957). *Admission to Grammar Schools*. London : Newnes.

THE SPEECH AND LANGUAGE DEVELOPMENT OF 5-YEAR-OLD CHILDREN

BY OLIVE C. SAMPSON

(City of Leicester School Psychological Service.)

SUMMARY. 1.—The aim of this research was to continue the study of speech development reported in 1956. The same fifty children (twenty-five boys and twenty-five girls), representative of all elements in the community, were followed up at the age of 5 years. They were given Intelligence, Vocabulary and Language tests. Most were seen in school.

2.—Every child had made progress and this progress appears to be related to his family's occupational status and his own general intelligence.

3.—When the earlier and present ratings were compared, the correlation was positive but somewhat low.

4.—There was some evidence in the performances of the 5-year-old children tending to indicate the boys' superiority in precision and the girls' in fluency of expression, but on the whole the sex differences were small.

I.—INTRODUCTION.

THE following is a continuation of " A Study of Speech Development of Children of 18-30 months " which was reported in this *Journal* in 1956.[14] The same fifty subjects (twenty-five boys and twenty-five girls), constituting a representative sample of the population according to their father's occupational status, were again seen. The object on this occasion was to continue the longitudinal study which had been started so as (1) to obtain more information about the continuity and ramifications of speech development, (2) to throw further light on the significance of the earlier findings.

II.—DESCRIPTION OF METHOD.

Whereas previously the children had (with very few exceptions) been seen at home, on this occasion the majority (forty-three) were interviewed at school, though contact was made with the parents by letter. This arrangement brought forcibly to light the mobility of the families. Twenty were found to have moved house, two having left Leicester, while a third had been to Australia and back between 1956 and 1958. School visits involved twenty-two different schools. The seven exceptions who had to be seen at home were three girls and four boys, and they included the four private school pupils of the sample. The aim was to see each child as close to his or her fifth birthday as possible, but owing to the incidence of childish ailments and other causes, this arrangement could not be adhered to perfectly. The following table shows the actual ages at which the interviews occurred.

TABLE 1

AGES OF CHILDREN.

Age	4·11	5·0	5·1	5·2	5·3	5·4	5·5	Total
Boys	4	9	7	3	2			25
Girls	1	15	6	2			1	25
Together	5	24	13	5	2		1	50

In the course of each interview (which was divided if this seemed necessary), the child was given the following tests :

(1) Goodenough Drawing-of-a-Man.[10]
(2) Revised Stanford-Binet Intelligence Scale, Form L.[16]
(3) Watts' Vocabulary Test for Young Children.[18]
(4) Watts' English Language Scale.[18]
(5) The children wrote their names if they could.

All the testing was done by the writer, but the Goodenough drawings were scored by another psychologist.

III.—TABULATION OF RESULTS.

(*a*) *The Intelligence Tests.*

TABLE 2

GOODENOUGH DRAWING-OF-A-MAN TEST.

I.Q.	Boys (19)	Girls (24)	Both (43)
Mean	104	111	108
Median	108	114	108
s.d.	11·9	16·9	15·6

TABLE 3

TERMAN-MERRILL SCALE (FORM L).

I.Q.	Boys (25)	Girls (25)	Both (50)
Mean	110	113	111
Median	113	113	113
s.d.	12·5	14·6	13·8

TABLE 4

CORRELATION BETWEEN THE INTELLIGENCE TESTS.

	Boys (19)	Girls (24)	Both (43)
Correlation	+·35	+·48	+·41
p.e.	±·138	±·141	±·079

(*b*) *The Vocabulary and Language Tests.*

TABLE 5

THE VOCABULARY (100 WORD) TEST.

Scores	Boys (25)	Girls (25)	Both (50)
Mean	33	30	32
Median	38	30	32
s.d.	9·7	9·6	9·9

TABLE 6

THE LANGUAGE TEST.

Scores	Boys (25)	Girls (25)	Both (50)
Mean	35	37	36
Median	36	39	37
s.d.	10·5	10·8	10·9

TABLE 7

INTERCORRELATIONS BETWEEN MENTAL AGE, VOCABULARY, LANGUAGE AND Q RATING.

	Language	Vocabulary	Q
M.A.	+·74(±·043)	+·73(±·044)	+·34(±·084)
Language		+·69(±·049)	+·43(±·077)
Vocabulary			+·42(±·078)

TABLE 8

AVERAGE SCORES AND OCCUPATIONAL STATUS.

Status	N	Q	I.Q.	Vocabulary	Language
A and B (Professional)	9	48	120	37	42
C (Skilled)	26	44	112	33	38
D and E (Unskilled)	15	29	104	26	27

TABLE 9

NAME WRITING.

	Boys	Girls	Both
Successful Attempts	4	10	14
Unsuccessful attempts	4	4	8
Unable to attempt	17	11	28

IV.—DISCUSSION.

(*a*) *The Intelligence Tests.*—Intelligence tests were introduced so as to provide, as far as possible, an objective measure of the children's mental development. It was hoped that this would throw light on the meaning of their progress in speech and language. Testing at this age is, of course, full of uncertainites,[17] and in so small a sample these undoubtedly affected the overall result considerably.

The Drawing-of-a-Man test gave the Group a mean I.Q. of 108 (s.d. 15·6) and the correlation between the rankings on this and on the Stanford-Binet Scale were low ($\rho = \cdot 41$), and not reliable. In the six cases of most marked discrepancy only one was placed lower on the Goodenough than on the Terman-Merrill Scale. In some cases poor ' Man ' results seemed to reflect maladjustment rather than low intelligence, and as the test was refused by six boys and omitted with one girl, no attempt was made to correlate it with the vocabulary and language results.

The Terman-Merrill Scale gave a mean I.Q. of 111 (s.d. 13·8) for the group, the median being at I.Q. 113. This would suggest that the sample as a whole is considerably above average, were it not known that, particularly for the age in question, many of the test items are somewhat too easy, so that the group's appearance of superiority may be an artifact of the test. Evidence in support of this can be found in articles by Burt[3], Kennedy Fraser[7] and Baldwin[1, 2]. A particularly interesting parallel to the present findings is afforded by the results of ninety-four urban New Zealand children of mean age 4½, who were likewise a representative sample selected according to the father's occupational status.[6] These obtained I.Q. 113 (s.d. 15·5).

An analysis of the test responses was made and this brought out the ' lack of variability ' on which Goodenough[9] remarks as " characteristic of about the age of 6 years."

TABLE 10

PASSES AT EACH LEVEL ON TERMAN-MERRILL SCALE.

Mental Year	No. of Cases	Possible Successes	Actual Successes
III	1	6	5
III—6	3	18	12
IV	13	78	73
IV—6	22	132	102
V	45	270	206
VI	49	294	200
VII	39	234	23
VIII	7	42	10
IX	1	6	2
X	1	6	0

A striking point which emerges is that the tests for year V and VI seem of too nearly the same difficulty for these children just turned five. At year V, 76 per cent. passes and at year VI, 67 per cent. passes are obtained. In contrast year VII presents a steep increase in difficulty and only 10 per cent. of passes occur.

In spite of these defects the Terman-Merrill provided an independent and objective order of merit and this served the purpose for which it was given.

(*b*) *The Vocabulary and Language Tests.*—At the interviews given when the children were 18–30 months of age, their language skill was analysed on the basis of verbatim reporting of free conversation on the lines suggested by McCarthy's work[12]. This method has been recently used with children up to 8 years of age by Templin,[15] but as an approach on more structured lines was felt to be desirable, the " Vocabulary Test for Young Children " and the " English Language Scale " described by Watts in " The Language and Mental Development of Children " were preferred.

The Vocabulary Test consists of 100 questions to be answered orally and individually. They are grouped in sets of five according to subject matter and difficulty and are designed to produce single word responses, e.g., Q : " What part of my face am I touching?" A: "Nose." Watts quotes figures to indicate norms for children from 3½ to 8½ years of age. With the present group very exact standards of correctness were required and errors were noted. The resulting average scores were considerably lower than those quoted by Watts. The errors were of four kinds. They consisted of descriptive circumlocutions ('nose holes' for 'nostrils,' 'hair by eyes' for 'eyebrows'), functional definitions ('for when you sneeze' for 'nostrils,' 'going to thump' for 'clenching') or interpretations. These sometimes suggested the autistic thinking that Piaget[13] observed in his young subjects, e.g., a square drawn by the examiner was named a 'window,' 'box' or 'house,' a triangle was a 'tent' or 'cheese.' Lack of knowledge accounted for another group of errors, as when a half-crown piece was called '£100'!

The language test consists of a series of sets of pictures for description by children from 4 to 10 years of age. The examiner describes the first two pictures in each set and lets the child imitate him. The child then proceeds to describe the other pictures in the set for himself and is scored on his performances. The situations pictured and the grammatical formulae required increase in complexity with each set.

Owing to the age of the children and the desire to explore as far as possible their ability to express themselves, the scoring was modified so as to give credit for any answer which revealed clear understanding of the meaning of the pictures, even when the grammatical formula expected was beyond the child. Thus, while full marks (3) were given (in accordance with the rubric) for such an answer as "That little boy's throwing a fish to see if the cat can catch it," "boy throwing a fish, cat trying to get it," got some credit (1), though "boy looking at a fish and a cat jumping" did not score.

Intercorrelations between these tests, the Terman-Merrill Scale, and the original terminal rating (known as the Q score) at 2½ years threw light on the children's speech development as follows:

(i) *Intercorrelations of the tests at* 5 *years.*—Considering the smallness of the sample and the variability of performance to be expected at this early age, the intercorrelations between the Intelligence, Language and Vocabulary tests indicate a considerable amount of agreement.

(ii) *Intercorrelations with* Q.—As the present investigation is part of a longitudinal study, the correlations between the order of merit originally established (Q Rating) and the order now obtained, is of crucial interest. The low, but positive, correlation obtained suggests the need for caution in making predictions on the basis of early performances. Study of the upper and lower quartiles shows that this was particlarly true in the case of poor early results, which were often improved upon, whereas early high scoring tended to be maintained. Worthy of note, also, was the fact that the children's Q performances are more in line with their 5-year-old ratings in Language and Vocabulary than they are with the Intelligence test results, pointing to a genuine continuity in the specific speech skill.

(iii) *Speech and Occupational Status.*—A very coherent developmental picture occurred when the results were studied according to the occupational status of the fathers. Descoeudres[5], noting the same phenonomenon, believes that the differences in achievement are related to "la condition materielle" rather than to "le niveau intellectuel" of the families. But the two are far from being separable and the present writer would be inclined to give greater weight to the intellectual factor.

(iv) *Boys and Girls.*—On the whole the differences between the boys' and girls' scores are not great, but it is somewhat striking that the boys, though on the average they are of lower intelligence, obtain higher scores on the Vocabulary test than do the girls. This, and the superiority of the girls on the Language test, confirm Templin's[15] findings and seem to endorse the advantages traditionally predicated for the sexes in precision and fluency respectively. In this connection, it is interesting that even in 1956 the results with this sample, which in general showed a ' striking equality ' between the boys and girls, nevertheless gave the girls a slight lead in total vocabulary. At 5, the girls also show superiority in name writing and the incidence of left-handedness is, as expected,[4] less, one girl and three boys being noted as left-handed. In none of these cases, however, was left-handedness related to any weakness in speech skills.

(v) *Correctness of Speech.*—No attempt was made to investigate this in detail. On the whole, the standard of speech seemed to the examiner satisfactory in that it was generally clear and confident, the sort of language inhibition which Gesell[8] notes being almost entirely absent. Childish errors, of course, appeared but only in five cases (three boys and two girls) did the faults seem to be serious : one of the girls had a very nasal tone and was said to be due for tonsillectomy, another (previously liable to temper tantrums) had an unduly timid production; of the boys the poor enunciation of one seemed probably due to his background, while the marked dyslalia and slight stammer of two others seemed connected with their personal maladjustment.

V.—REFERENCES.

[1] BALDWIN, A. L. (1948). The relative difficulty of Stanford-Binet items and their relation to I.Q. *Journal of Personality*, 16, 417-429.

[2] BALDWIN, A. L. (1948). Variations in Stanford-Binet I.Q. resulting from an artifact of the test. *Journal of Personality*, 17, 186-198.

[3] BURT, C. (1939). The latest revision of the Binet intelligence tests. *Eugenics Review*, 30, 255-260.

[4] CLARK, M. M. (1957). *Left-handedness*. University of London Press.

[5] DESCOEUDRES, A. (1946). *Le Developpement de L'Enfant*. Delachaux et Niestle.

[6] FITT, A. B. (1952). *The Stanford-Binet Scale : Its Suitability for New Zealand*. N.Z. Council for Educational Research.

[7] FRASER, D. K. (1948). *The Terman-Merrill Intelligence Scale in Scotland*. University of London Press.

[8] GESELL, A. (1940). *The First Five Years of Life*. Methuen.

[9] GOODENOUGH, F. L. (1942). Variability of the I.Q. at successive age levels. *Journal of Educ. Psychol.*, 33, 241-251.

[10] GOODENEOUGH, F. L. (1926). *The Measurement of Intelligence by Drawings*. World Book Co.

[11] MCCARTHY, D. A. (1946). Language development in children. *Manual of Child Psychology*. Chapman and Hall.

[12] MCCARTHY, D. A. (1930). *The Language Development of the Pre-School Child*. University of Minnesota Press.

[13] PIAGET, J. (1948). *The Language and Thought of the Child*. Routledge.

[14] SAMPSON, O. C. (1956). A study of speech development in children of 18-30 months. *Brit. J. Educ. Psychol.*, 3, 194-201.

[15] TEMPLIN, M. C. (1957). *Certain Language Skills in Children*. University of Minnesota Press.

[16] TERMAN, L. M., and MERRILL, M. A. (1939). *Measuring Intelligence*. Harrap.

[17] UPDEGRAFF, R. (1932). The determination of a reliable intelligence quotient for the young child. *Journal Gen. Psychol.*, 41, 152-166.

[18] WATTS, A. F. (1944). *The Language and Mental Development of Children*. Harrap.

(*Manuscript received, September 20th*, 1958.)

THE CLASSIFICATION OF INSTANCES OF FOUR COMMON CLASS CONCEPTS BY CHILDREN AND ADULTS*

By MARIAN ANNETT
(*Warneford Hospital, Oxford.*)

SUMMARY. Dispositional views of conceptual thinking suggest the need for a descriptive study of the ways in which normal subjects of various ages classify familiar objects. The sortings and explanations of 303 children, aged 5-11 years, and forty-two adults, aged 18-73 years, in grouping pictures of common objects, were analysed. The explanations were classified for both form and content. Five methods of explanation were distinguished, only two of which are recognised in logic. There were four main types of content, the most numerous at all ages being those concerned with spatial position and activity. The main developmental changes were found in methods of explanation. Older adults perform more like children than younger adults.

I.—INTRODUCTION.

THE need for a study of methods of classification and the way in which it should be undertaken was suggested by dispositional views of conceptual thinking (such as those of Wolters 1933, Piaget 1950 and Price 1953). If we agree that concepts show themselves in dispositions to behave appropriately to recurring features of the environment and agree also that most, if not all, concepts are acquired gradually through experience, then we can expect concepts to be manifested during early stages of growth in more or less inappropriate ways. The dispositional view suggests that, in studies of the growth of concepts, it is not enough to look for the earliest behaviour which the experimenter recognises as correct and treat all responses which fall short of this criterion as wrong or concrete without further description. Piaget has shown, in numerous experiments, that children make systematic mistakes before they perform in the manner expected by adults and that study of such mistakes is essential in tracing the development of a concept.

The present experiment is concerned to trace some aspects of the growth of some common class concepts. Concepts involve the classification of experiences, and methods of classification may well change during the course of ontogenetic growth. To investigate such changes it is necessary to devise tasks which can be given in a standard manner to subjects varying widely in age and intelligence, from the youngest children who can perform the task to intelligent adults (for it may be as unwise to make assumptions about the final stages of growth as about the earliest). Methods of classification are probably determined by a large number of factors such as the subjects' familiarity with the items, the range of items and, in particular, by the subjects' purpose in making the classification. No single study could be expected to elicit examples of all possible methods of classification, nor would the frequencies of methods at various ages hold for tasks of different complexity. The purpose of the present study was first, to discover how normal subjects most readily classify a group of sixteen common objects when they are given as little direction as possible about the way in which they are to proceed and secondly, how methods of classification change with age and intelligence. A subsidiary experiment was made to investigate the influence of instruction on certain aspects of the results.

---

* Based on a paper read at a conference of the British Psychological Society in December, 1957.

II.—The Investigation.

(1) *Material.*—The material consisted of 16 cards on each of which was an uncoloured line drawing. Four of the cards portrayed animals (cow, bird, fish, butterfly), four plants (tree, flower, apple, toadstool), four vehicles (car, train, ship and aeroplane), and four articles of furniture (chair, desk, clock and television). The drawings were immediately recognised by the subjects although the toadstool was sometimes called a mushroom and the television a radio.

(2) *Procedure.*—Subjects were examined individually in a separate room or a secluded part of the school or ward.

They were first given a vocabulary test in order to make a rough assessment of intelligence, children being given the four vocabularies used by Dunsdon and Roberts (1957), and adults the Mill Hill A and B scales (Raven, 1943).

The cards were then laid before the subject, always in the same mixed order. Subjects were instructed as follows: " Some of these cards belong together. Sort them into groups so that the ones that go together are in the same group." Each of the subject's groups was laid before him in turn and he was asked, " How do these belong together? " The sortings and explanations were recorded in full.

(3) *Subjects.*—The 303 children aged 5-11 years were pupils at infant and junior schools in the City of Bath. Half the local authority schools, the E.S.N. school and three independent schools were visited, the proportion of children drawn from each source being L.E.A. 92·4 per cent., E.S.N. 1·3 per cent., and Ind. 6·2 per cent. All the children had had a birthday within the month preceding the examination, and all or a random selection of the children fulfilling this requirement in each school were seen. Age, sex and vocabulary means and standard deviation (using norms found by Dunsdon and Roberts, 1957) are given in Table 1.

TABLE 1

Numbers of Children at Different Ages, and Mean Vocabulary Quotients.

Age	5	6	7	8	9	10	11	Total
Boys	22	19	21	28	28	24	18	160
Girls	19	17	23	26	25	14	19	143
Total	41	36	44	54	53	38	37	303
Vocab. mean ..	106·5	95·0	100·3	97·2	105·3	104·6	96·6	100·9
S.D.	12·3	21·8	15·4	14·6	14·5	11·16	22·3	16·5

The 42 adults aged 18-73 years were patients in an Oxford orthopaedic hospital who volunteered after an introductory letter had been distributed in the wards. All could move their hands freely enough to inspect and sort the cards. They have been divided into two age groups, less than 40 and 40 plus, and into three vocabulary grades, according to combined Mill Hill A and B scores. Table 2 gives the age and vocabulary distribution.

TABLE 2

THE AGE AND VOCABULARY DISTRIBUTION OF THE ADULTS.

Age		Mill Hill Vocabulary Grades			
Range	Mean	IV	III	II and I	Total
18–39	27·04	9	11	3	23
40–73	55·15	3	10	6	19
		12	21	9	42

(4) *Results.*

A.—*Sortings.*—A complete 4 by 4 sorting into animal, plant, vehicle and furniture groups was made by only twenty-four children and eleven adults, and only one child and one adult made a complete 2 by 8 sorting into living and non-living things. There was wide variability at all ages in the number of groups made and the number of cards in a group. The most frequent response at years five and six was to make no sorting at all, at seven and eight to sort the cards in eight pairs and at nine and over, to make four or five groups. About 50 per cent. of the children under nine years grouped the cards only in pairs.

More than 200 different combinations of the cards were produced, but half of these occurred only once. Some of the most frequent sortings were fish and ship, bird and tree, butterfly and flower, and combinations of bird, butterfly and plane; these sortings occur in all age groups. Fig. 1 shows the percent. subjects in age groups who sorted animals with plants, vehicles with furniture and living with non-living things (apart from the four sortings mentioned above), together with subjects sorting within the four main groups (that is, combinations of animal cards only, and similarly plants, vehicles and furniture) and subjects making no sorting. All but a few younger children made some sorting. Instances of all types of sorting occur in all age groups but there is a clear rise of within group sortings and a decline of cross group sortings as age increases.* An exception to this decline is the rise of animal with plant groupings in 11-year-olds and adults. (More is said of this in section G below.)

B.—*Explanations.*—The classification below describes the explanations taking into account the form and content of the subjects' statements without reference to the question of whether they were correct. It is possible to distinguish clearly between five types of statement which will be called methods of explanation (main headings) and between various types of content (sub-headings in Nos. 2, 3, and 4.). The explanations of all but two children were readily classified according to this scheme; one exception was a sorting according to the first letter of the names of the objects and the second was an analogy.

Scheme of Classification.

(1) *No Explanation.*—This group includes no response, 'Don't know,' and assertions that the objects belong together without explaining how. For example, 'Just did' and 'They go together.'

(2) *Enumeration.*—These explanations give facts about each object in turn, but different facts about each one. The statements do not explain how the objects belong together; but they are quite distinct from those called 'No

---

* As subjects may have arranged the cards in from 1-8 groups and as one subject may have given more than one type of sorting or explanation, all frequencies are a percentage of subjects giving at least one of the responses under consideration.

FIG. 1

Sortings Classified.

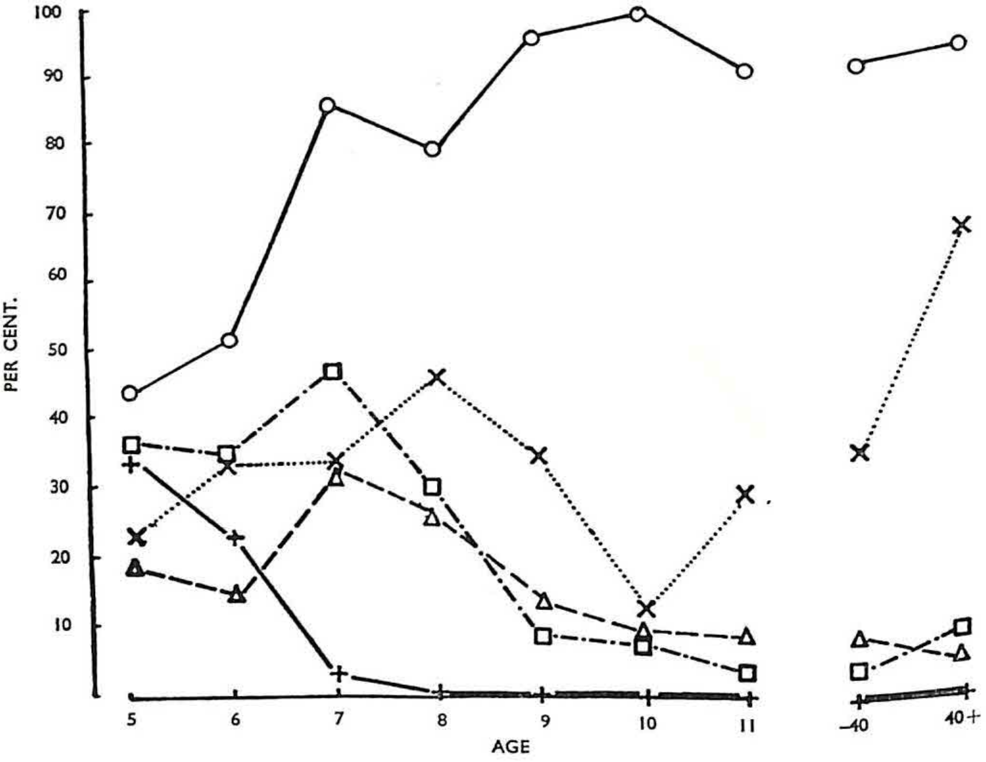


FIG. 1

Sortings Classified. Per cent. subjects giving at least one of each of the following sortings: Within Category (O———O), Animal with Plant (X·······X), Vehicle with Furniture (Δ- - - - - - -Δ), Living with Non-living (□—·—·—□), and No Sorting (+———+).

Explanation' in that something is said about the objects. Only one child directly compared the objects; all others gave differences in this enumerated or juxtaposed form. Enumerations are sub-divided as follows: *

(i) *One Object.*—S. describes only one of the objects he has grouped together, e.g., the tree, cow and flower are explained with, ' The tree is in the field.'

(ii) *Naming.*—Only the names of the objects are given, e.g., ' This is a cow, this is a bird and this is a tree.'

(iii) *Place*, e.g., ' The cow is on the farm and the fish is in the sea.'

(iv) *Activity*, e.g., ' The boat sails and the car runs.'

(v) *Structure*, e.g., ' The cow has a tail and the butterfly has wings.'

* Statements about similar qualities or class names can also be enumerated; it might be said that, 'The tree grows the flower grows and the apple grows,' or, 'This is plant and this is a plant,' instead of the shorter statements, ' These all grow,' and ' These are plants.' These enumerations have been included with other similarity and class name statements; the present section includes only those in which different facts are mentioned.

(3) *Contiguity.*—These explanations relate the objects together in a direct, concrete interaction. The objects are envisaged as impinging on one another in spatial contiguity. S's do not always make direct reference to place; they may stress the role of an animal or human being in bringing the objects together and sometimes a temporal relation is the important one. Examples of types of contiguity follow:

(i) *Place.*—'The apple grows on the tree.' 'Because the chair is in the plane.' 'The clock goes on the desk.'

(ii) *Time.*—'When the T.V. has been on for an hour you can see by the clock.' 'If you want to see T.V. you've got to know the right time.'

(iii) *Animal Activity.*—'Butterflies settle on flowers.' 'Cows eat mushrooms.' 'Birds build their nests in trees.'

(iv) *Human Activity.*—'The engine driver might be eating an apple when he's driving the train, or the fireman.' 'You sit in a chair to watch T.V.'

(v) *Serial Contiguities.*—A large number of objects may be related together in a series of contiguities, e.g., 'The farmer can tell the time and sit on the chair and keep things in the drawers and he has a T.V. to look at . . .'

(4) *Similarities.*—A characteristic which is common to the objects is given. The characteristics may be grouped as follows:

(i) *Place.*—'The fish and ship are in the sea.' 'The clock, desk and chair go in the house.'

(ii) *Activity.*—This includes animal, plant and mechanical activity, 'They fly,' 'They grow,' 'They go along.'

(iii) *Human Activity.*—'People ride in them,' 'You sit on them.'

(iv) *Structure.*—This may be in terms of the material of which the objects are made, 'They are all wooden,' or parts they have in common, 'They have wheels,' or visible similarities in the drawings, 'These have got little stalks.'

(5) *Class Name.*—A class name is given, e.g., 'Creatures,' 'Machines,' 'Household things.' The names may not be entirely appropriate; the animals might be called 'Pets' and the vehicles 'Toys.'

C.—*The Reliability of the Scheme of Classification.*—Three judges, none of whom had been associated with the experiment in any way, were asked to classify a selection of the explanations (extracted from the list of all explanations at fixed intervals) according to the above scheme. Two judges agreed with the writer 90 per cent. and one 88 per cent. as to method of explanation, and agreement as to content varied between 73–83 per cent. There was complete agreement between the three judges and the writer for 83 per cent. of the explanations as to method and for 66 per cent. as to content.

D.—*Inter-relations of the Five Methods of Explanation.*—In view of the possibility that some may wish to regard types of explanation as indicative of developmental stages, it is necessary to ask if S's consistently use one method or if they use more than one method in explaining (1) a single sorting and (2) several sortings.

With regard to the single sorting, S's can use more than one type of explanation in a statement about one group of objects. For example, they might say, 'These are plants and the apple grows on the tree and the flower grows under the tree' combining class name and contiguity; or, 'These all move, this on land, this in the sea and this in the air,' combining similarity and enumeration. In order to avoid judgments about which was the main or best explanation, both parts of these multiple explanations are given equal weight in the frequency counts. However, additional enumerations, like the second example above give rise to a special problem as they occur frequently, particularly in older adults.

50 per cent. of the older adults added these redundant, separate details to their class name or similarity statements, but only 5 per cent. gave enumeration alone. It would be misleadng to group the main and additional enumerations together and the graphs below show the frequency of enumeration as the main explanation.

In accounting for their several sortings, two-thirds of the children and more than half the adults change their method of explanation. Of the children who gave class names, for example, almost all gave similarities and more than half also gave contiguities. The children who gave only one type of explanation, on the other hand, consisted mainly of those giving no explanation (41 per cent.) and those giving contiguities (31 per cent.). Table 3 shows the association between the various methods in the children's explanations (Product Moment correlations). Many of the correlations are significant but none are so high as to suggest that the categories of explanation are not usefully distinguished.

TABLE 3

CORRELATIONS BETWEEN THE FIVE METHODS OF EXPLANATION.

	Enumeration	Contiguity	Similarity	Class Name
No Explanation	—·047	—·281*	—·404*	—·327*
S.E.	·057	·052	·048	·061
Enumeration		·153*	—·011	—·192*
S.E.		·056	·057	·055
Contiguity			·063	—·142
S.E.			·057	·056
Similarity				·345*
S.E.				·051

* P< ·01

E.—*Method of Explanation with Age and Vocabulary Score.*—As a comparison of the boys and girls explanations showed no significant differences, the sexes are combined in the following calculations.

Age.—The percentage of S's in each age group giving at least one example of each method of explanation is shown in Fig. 2. The frequency of no explanation falls sharply as age increases. Contiguities rise and fall with a peak at year 8 in children, and rise again in older adults.

Similarities show a roughly linear rise with age, preceding the growth of class name. Similarities are less frequent in adults, when class names distinctly supersede them. Enumeration as the only explanation of a sorting declines after year 7. If the additional enumerations had been included with the primary ones, they would be seen to remain between 20–30 per cent. in children and rise to 50 per cent. in older adults.

The bell-shaped distribution of the contiguities and the rise of contiguities and additional enumerations in older adults led to a re-examination of these answers, in case some difference between those of younger and older S's had escaped attention on first analysis. No difference between the contiguities of the older and younger S's was apparent on re-inspection. The adult enumerations, on the other hand, were found to consist largely of naming, e.g., 'These are

FIG. 2
Method of Explanation with Age

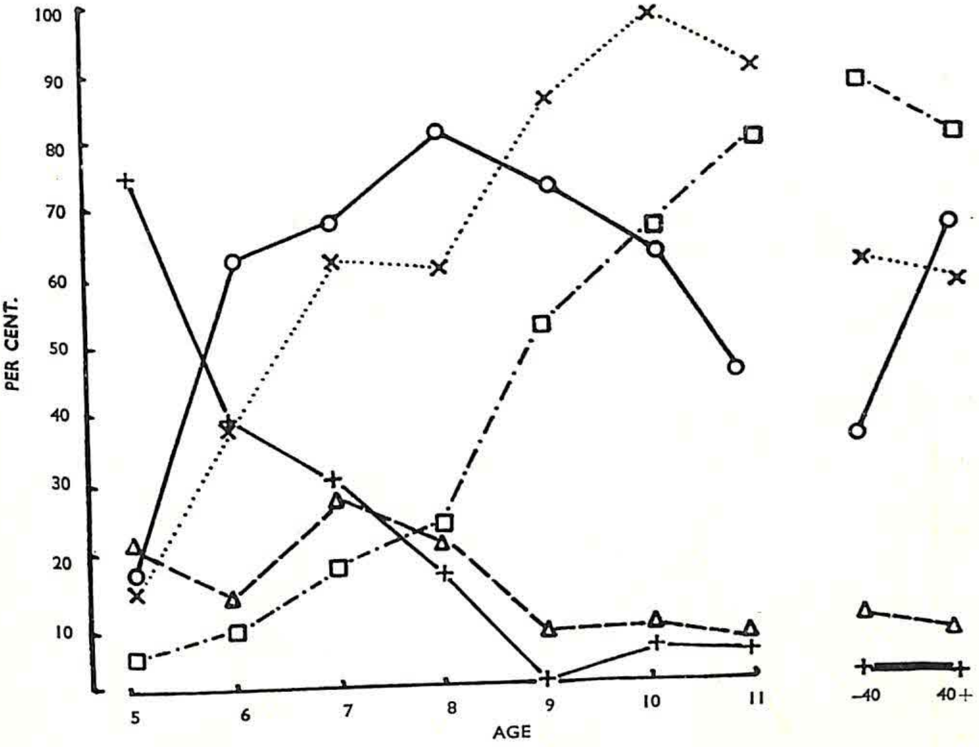


FIG. 2

Method of Explanation with Age. Per cent. subjects giving at least one example of each method of explanation: No Explanation (+———+), Enumeration (Δ------Δ), Contiguity (O———O), Similarity (X·······X), and Class Name (□-·-·-□).

plants, tree, flower and mushroom.' 89 per cent. of the adult enumerations were of this type in contrast with 22 per cent. in children.

Some of the youngest children give class names. This is not due to the fact that names such as 'Toys' and 'Pets' were included in this group: the 5 and 6-year-olds used the words 'Animal' and 'Furniture.'

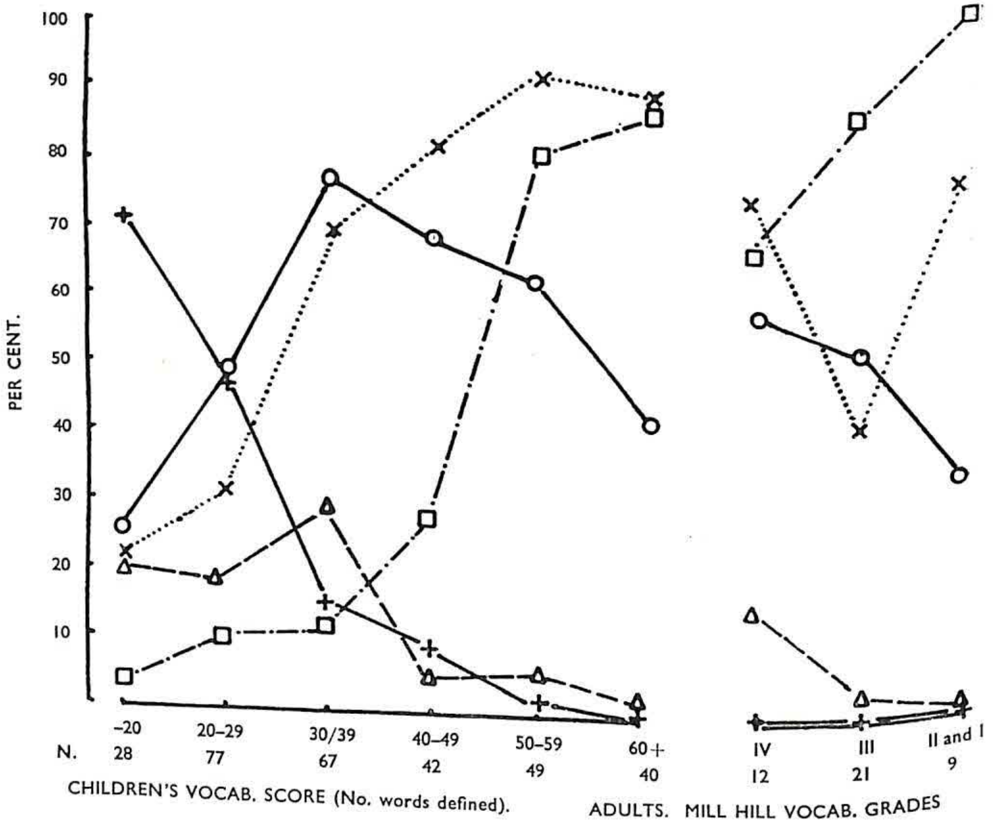
Vocabulary score.—The children are grouped here according to the total number of words defined correctly on the four vocabularies and the adults according to Mill Hill vocabulary grades. The percentage of S's giving each method of explanation is shown in Fig. 3. The distribution is essentially the same as that for age. Contiguities clearly precede similarities which in turn precede class names. Class names are given by children with very poor vocabularies.

All adults of above average intelligence gave at least one class name. In evaluating the adult distribution, it should be recalled that the intelligent adults tended to be older than those of lesser intelligence. The three above average adults who gave contiguities were 57, 65 and 73 years of age. The rise of similarity explanations from average to above average adults is due to the tendency of the intelligent (older) adults to make additional similarity statements (like additional enumerations) after the class name.

F.—*Content of Explanation with Age and Vocabulary Score.*—In enumerations, contiguities and similarities, S's are concerned with facts about place, activity and structure. The frequency of each of these types of content has been calculated without reference to the method of explanation involved. Time was mentioned by only three S's and is disregarded in the analysis below. Human activity has been calculated separately from animal, plant and mechanical activity in view of its bearing on the frequent observation that children tend to define and classify according to use.

FIG. 3

Method of Explanation with Vocabulary Score. (As Fig. 2.)



Age.—Fig. 4 gives the percentage of S's in age groups who make at least one explanation involving content (that is, all S's who did not consistently give no explanation or class name) and the percentage giving each type of content. The use of each type of content, in relation to total content, is roughly constant at all ages. Place is slightly more frequent than activity in most groups, but the differences are not significant. Human activity and structure are consistently less frequent than place and activity and have a similar distribution.

FIG. 4
Attributes with Age

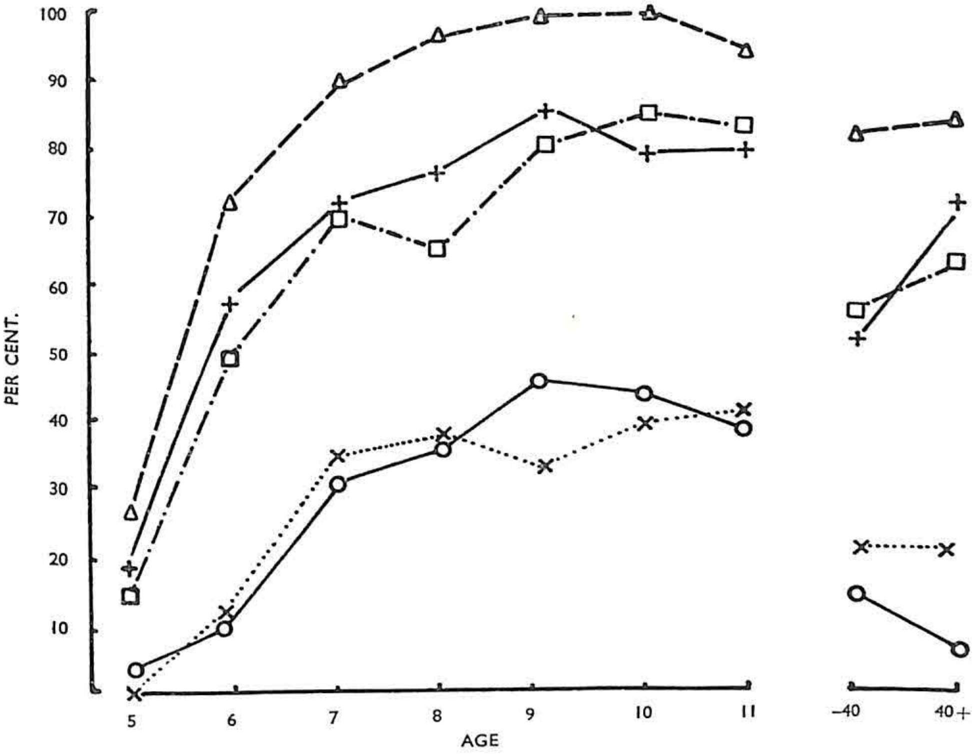


FIG. 4

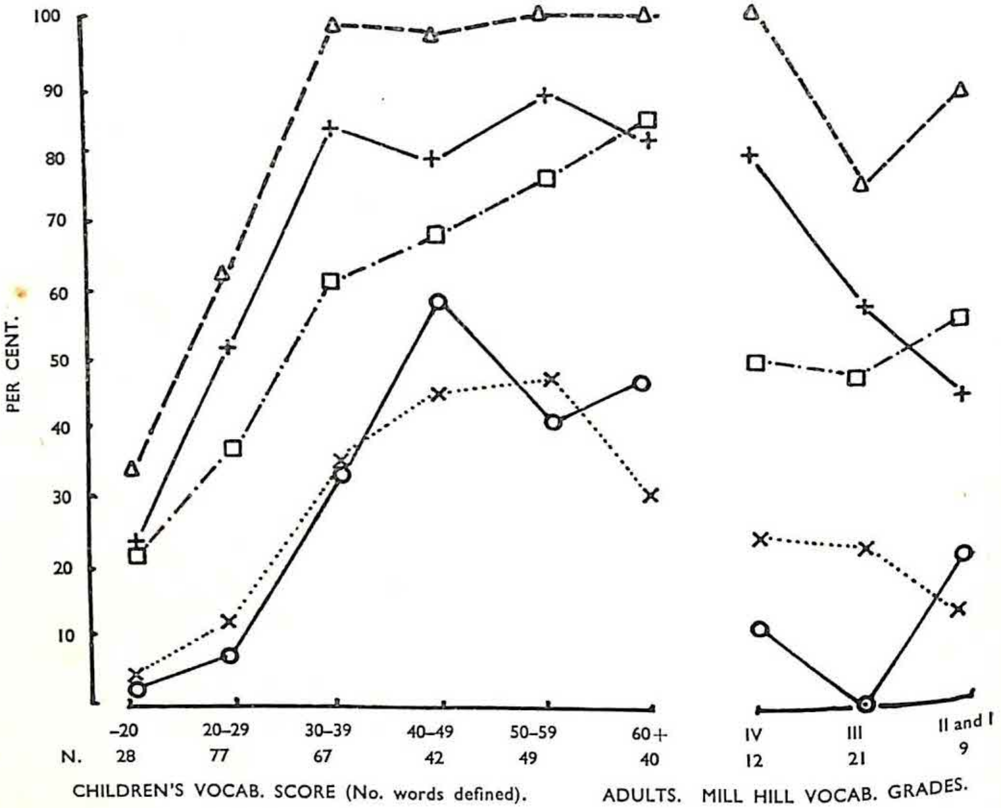
Content with age. Per cent. subjects giving at least one explanation involving content (Δ- - - - - - Δ) and those giving at least one of Place (+———+), Activity (□—·—·—□), Human Activity (×······×), and Structure (o———o).

Vocabulary score.—The distribution of content in vocabulary score groups is like that in age groups. Adults of average and higher intelligence mention fewer facts about the objects as they more often given a single class name statement. The proportions of types of content remain more or less constant, apart from a small change in intelligent children and adults which suggests that activity explanations are superseding those of place (see Fig. 5).

G.—*The Relation of Explanation to Sorting.*—Sortings including cards from more than one of the four main groups tend to be explained in a characteristic manner at whatever age they appear. Fish and ship and combinations of things that fly are usually explained through similarity. All other cross category sortings are usually explained through contiguity. This holds true even of animal with plant sortings, which were seen in Fig. 1 to be common in 11-year-olds and adults. Of the twenty-one adults who sorted in this way, only four said the pictures were of living things and thirteen gave a contiguity explanation.

FIG. 5

Content with vocabulary score. (As Fig. 4.)



Sortings within the four main animal, plant, vehicle and furniture groups may be explained by any of the five methods but they tend to be associated with particular methods. Animal sortings are most often explained by class name and secondly by similarity in all age groups. Vehicle sortings are mainly explained through similarity; class name does not supersede similarity until adulthood. Furniture sortings are commonly explained through contiguity except in 11-year-olds and younger adults; older adults return to contiguity. Plant sortings are explained equally by contiguity, similarity and class name. In all these sorting groups, the use of similarity and class name increases with age until class name is more frequent than all other methods of explanation (except for furniture sortings of older adults as mentioned above).

III.—SUBSIDIARY INVESTIGATION: THE INFLUENCE OF INSTRUCTION ON METHOD OF EXPLANATION.

It might be supposed that some explanations, particularly contiguities arise from the instruction, 'Put those that *belong* together into groups'; the use of 'belong' may suggest groupings of the type cow and mushroom, 'Because cows eat mushrooms,' and clock and desk, 'Because the clock is on the desk.' A group of Oxford children, comparable with the Bath 8-year-olds in age and vocabulary score, were examined as in the main investigation but substituting the instruction, 'Put those which are *alike* together.' Table 4 compares the two groups of children for age and vocabulary score and the number of children in each group giving each of the five methods of explanation. Of the children instructed to put those 'alike' together, more gave similarity explanations than those instructed with 'belong,' but there was no significant change in the number giving contiguities or any of the other three methods of explanation. It seems, therefore, that the frequency of contiguity in the main investigation is not due to the use of 'belong' in the instruction.

TABLE 4

A COMPARISON OF TWO GROUPS OF SUBJECTS GIVEN DIFFERENT INSTRUCTIONS FOR AGE AND VOCABULARY SCORE (USING t), AND FOR NUMBERS OF CHILDREN GIVING EACH OF FIVE METHODS OF EXPLANATION (USING χ^2 WITH YATES' CORRECTION FOR CONTINUITY)

Instruction	Age (months)	Vocabulary (no. of words)	No Expl.	Enum.	Contig.	Simil.	Class N.
	Means		Number of Children				
'Belong' N 54	96·0	36·3	9	12	43	34	13
'Alike' N 31	97·5	35·3	11	6	21	27	13
	t=1·91	0·44	χ^2=2·90	0·00	0·93	4·50*	0·22
			1 d.f. *P<·05.				

IV.—CONCLUSIONS.

The objects used in this study were classified in five different ways. In the *first*, S either does not attempt to explain or simply asserts that the objects belong together. No facts about the objects are mentioned. The *second* sort of statement (enumeration) rests entirely on the description of characteristics. Certain qualities are found noteworthy about each object and S describes them in turn, without reference to the qualities of other objects grouped with them. These statements are a catalogue of different facts joined together by the conjunction 'and.' Terman and Merrill (1937) among others have found that differences are often given in a test of similarities. Piaget (1928) found a similar method of explanation, which he called Juxtaposition, when studying children's use of conjunctions of causality. In contrast with the first type of explanation above, it seems here that S analyses significant features from the total complex of attributes presented by the object, ; but he does not consider the attributes of two objects at once. Children and older adults often use statements of this kind as a tail-piece to a statement of similarity or class name ; it is as though they need to remind themselves of the distinctions between objects they have grouped together.

The *third* type of explanation depends on physical contiguity in space. It describes the relations which actually obtain between objects ; S is concerned with how, in fact, the world is organised and what goes with what in everyday experience. Contiguity explanations are numerous in children and not infrequent in adults. They are the main precursor of similarities and class names. As far as the writer is aware, they have not been separated from other statements in previous studies of children's explanations. Thompson (1941) has a category which includes groupings 'according to a concrete situation,' among others. Reichard, Schneider and Rapaport's (1944) Concrete category appears to include answers here called contiguity and similarity of structure with other 'inadequate' (their term) responses. The bell-shaped distribution of contiguities found in the present study means that these answers could not be used to predict age or intelligence as required in the construction of most tests; this fact has probably contributed to their being overlooked.

The *fourth* and *fifth* types of explanation are the familiar similarity and class name statements. S gives a quality or name which is common to the objects. These five types of explanation and no others were also found when children were asked to put objects which are alike together. They cannot be attributed to the use of 'belong' in the instructions of the main investigation.

All but the first and last types of explanation involve the mention of facts about the objects. A limitation on the variety of facts used was necessarily imposed by the nature of the material; for instance, there was no opportunity to classify according to colour or size. The facts used in this study fall into four main groups; they are concerned with location, activities arising from the objects themselves, activities performed by human beings and structural properties. A distinction should be drawn between the use of location as a characteristic of an object and as a means of relating objects together. S's use location in both ways; in all contiguity explanations, objects are related in that they occur together in the same place. But in some contiguities, as in some enumerations and some similarities, S's treat location as a noteworthy characteristic of a particular object. Place is used as an attribute in the same way that activity and structure are used as attributes. Of the four types of facts mentioned, facts of place are most frequent. This finding, together with that of the frequency of contiguity explanations, demonstrates that location has an important role in human thinking. In comparison with place, structural properties, including shape, material and other perceptible features are mentioned infrequently; this finding points to the limited relevance of experiments based on the assumption that concepts are acquired through observation of physical qualities. Explanations through use or human activity are considerably fewer than those through activities performed by the objects themselves; this result is unexpected in view of the prominence given to Functional explanations in the classification of Reichard, *et al.* (1944).

Bearing in mind the limitations on methods of classification imposed by the material presented and the fact that associations between methods and age or intelligence are a function of the difficulty of the particular objects used, a number of conclusions can be drawn about the way in which groupings and explanations change or do not change with development.

(1) Sortings show a clear change in that the number of combinations is reduced as age increases. The younger children produced a very large variety of groupings, most of them cutting across the four main groups; few particular sortings were repeated by different children. As age increases, the diversity is reduced and S's tend to sort within the four main categories.

(2) The sorts of facts found noteworthy about the objects do not change with age. The order of frequency is first, place and activity and secondly, human activity and structure. The use of facts of all kinds increases with age in children and then falls slightly as older children and adults rely on the class name alone. The present findings suggest that although concepts may develop

through the acquisition of facts about objects, facts about where objects are to be found and what they do are always of greatest interest.

(3) The most significant change with growth concerns the way in which facts are used in relating objects together. They may be juxtaposed, they may be the basis of interaction in contiguities, they may be qualities seen to be common to a number of objects, or they may be made redundant by class name.

(4) The order of development is through no explanation, enumeration, contiguity, similarity and finally class name whether the data are analysed for age or intelligence. It seems that children first need to analyse the characteristics of individual objects and then to consider how objects are organised in space before they can attend to significant similarities. In this last stage, it is often important to recall distinctions between objects grouped together (in additional enumerations). Similarity and class name statements show an almost linear rise with age which is not unexpected in view of their long use in tests of intelligence.

(5) These five methods of explanation are the only features of the results which could be interpreted as representative of stages of development. It might be objected that one cannot talk of stages here as four methods occur in all age groups and only one-third of the subjects use one method consistently. Deutsche (1937) concluded from similar findings that there is no evidence that children's concepts of causal relations progress by stages. However, whether or not concepts develop through the same stages, they need not all be at the same stage simultaneously. This is apparent from a test such as the Revised Stanford Binet in which it is expected that some concepts will be achieved at 7 years and others not until adulthood. To suggest that stages should be characteristic of age groups or individuals is to suggest that all problems of conceptual thinking are equally difficult. It is worth noting that the distribution of explanations found in this study is remarkably like that found by Russell and Dennis (1939) in their study of Piaget's stages of animistic thinking.

(6) In adults, explanations acceptable for correct, logical thinking generally predominate but the less acceptable methods are not infrequent. Older adults more often give enumerations and contiguities than younger ones. Many adults, like the children, use more than one method of explanation, going from one type to another without any awareness of the change or that one method may be better than another.

These findings support the Dispositional view in its implication that concepts may be attained through stages which are not apparent on consideration of their final form in adults. Despite the fact that problems of conceptual cognition have concerned intelligent minds for centuries, the subjects' treatment of familiar material in this study could not have been predicted on any theoretical ground. Until more descriptions of response to everyday material have been made, it is unlikely that pertinent hypotheses about the growth of real concepts can be formulated.

ACKNOWLEDGEMENTS.—This work was begun at the Burden Mental Research Department, Bristol, by kind permission of Dr. J. A. Fraser Roberts. Thanks are due to Dr. M. I. Dunsdon for her interest and help in the initial stages of the work ; to the City of Bath Education Committee, the Director, Mr. Brand, to the City of Oxford Education Committee, the Director, Mr. Garne, and to Head Teachers for making children and testing facilities available ; to Dr. C. W. M. Whitty and Mr. J. C. Scott for their help in the provision of adult subjects ; to Mrs. Hermelin, Mrs. Mittler, and N. O'Connor for acting as judges ; and to Mr. J. F. Scott for help with statistics.

V.—REFERENCES.

DEUTSCHE, J. M. (1937). The development of children's concepts of causal relations. *Univ. Minnesota Inst. Child Welf. Monogr. Series*. Univ. Minnesota Press.

DUNSDON, M. I., and ROBERTS, J. A. F. (1957). A study of the performance of 2,000 children on four vocabulary tests II. *Brit. J. Stat. Psychol.*, 10, 1-16.

PIAGET, J. (1928). *Judgment and Reasoning in the Child.* Eng. Tr. London : Routledge and Kegan Paul.

PIAGET, J. (1950). *The Psychology of Intelligence.* Eng. Tr. London : Routledge and Kegan Paul.

PRICE, H. H. (1953). *Thinking and Experience.* London : Hutchinson.

RAVEN, J. C. (1943). *The Mill Hill Vocabulary Scale.* London : H. K. Lewis.

REICHARD, S., SCHNEIDER, M., and RAPAPORT, D. (1944). The development of Concept formation in children. *Amer. J. Orthopsychiat.*, 14, 156-161.

RUSSELL, R. W., and DENNIS, W. (1939). Studies in Animism, I-IV. *J. Genet. Psychol.*, Vols. 55-57.

TERMAN, L. M., and MERRILL, M. A. (1937). *Measuring Intelligence.* Boston : Houghton, Mifflin.

THOMPSON, JANE (1941). The ability of children of different grade levels to generalise on sorting tests. *J. of Psychol.*, II, 119-126.

WOLTERS, A. W. (1933). On conceptual thinking. *Brit. J. Psychol.*, 24, 133-143.

A FACTOR ANALYSIS OF THE WECHSLER INTELLIGENCE SCALE FOR CHILDREN

By A. E. MAXWELL

(*Institute of Psychiatry (Maudsley Hospital), University of London.*)

Summary. A factor analysis of the correlation matrix for the sub-tests of the WISC was carried out for each of the age groups 7½, 10½ and 13½ years. In each analysis two factors were extracted and these, after rotation, are found to give patterns of loadings which are similar for all three age groups. The factors were then rotated : the first is labelled ' vg ' (verbal-intellectual), and the second ' sp ' (space-performance).

Methods for obtaining factor scores are then given and it is argued that such scores are preferable to the cumstomary *Verbal* and *Performance* scores when summarising a child's performance on this battery of tests.

I.—Introduction.

When using the WISC it is customary to obtain composite *Verbal* and *Prrformance* scores from the two groups of sub-tests bearing these labels, and to use them as a summary of a child's performance on the test as a whole. Though these composite scores are correlated to the extent of 0·6 or more, there is a tendency to think of them as defining two relatively independent mental factors. Wechsler himself seems to imply this when he says (6, p. 5) that " the tests identified as verbal and performance differ as these labels indicate." But if it is thought desirable to divide the tests into sub-groups, each measuring a separate aspect of cognitive ability, it would appear preferable to proceed along factor analytic lines. In this way it is possible to discover just how many *independent* factors are necessary to account for the co-variation between the variables and, after suitable rotation, to label them in a meaningful way. In addition, we can—given a child's scores on the individual tests—estimate his scores on the factors we have derived. Such scores, unlike the composite and highly correlated Verbal and Performance scores, will be relatively independent and by virtue of this independence will give a much clearer and more economic description of the child's cognitive ability. To this end a factor analysis of the matrix of correlations between the sub-tests of the WISC (6, pp. 10-11) for each of the age groups 7½, 10½ and 13½ years is reported here. In addition, a set of weights for obtaining a child's factor scores when his test scores are known is obtained and their use for describing the child's cognitive ability is illustrated.

II.—The sample.

When reporting the correlations, Wechsler states that each of the three matrices is based on a sample of 100 boys and 100 girls, including rural and urban children roughly in proportion to the numbers of such children in the U.S. population. Each sample claims to be on the whole a fair stratified sample where occupation of fathers is concerned ; only white children, however, have been included.

III.—The procedure.

Excluding from the battery the sub-tests *Information* and *Mazes* (which are seldom used at this Institute), but including the composite Verbal and Performance variables, a centroid analysis of each of the three correlation matrices already referred to was performed and two factors extracted. The loadings thus obtained were used as first approximations to the 'true' factor loadings and maximum likelihood estimates of these, using Lawley's[1,2] method of analysis, were obtained. These appear in Table 1. The residuals, after two factors, were found to be so small (the correlations are reported to two decimal places only) that it was decided to be unwise to extract further factors.

TABLE 1

MAXIMUM LIKELIHOOD LOADINGS.

	I	II	h^2	I	II	h^2	I	II	h^2
	(7½ years)			(10½ years)			(13½ years)		
1. Comp.	·59	·13	·36	·76	·06	·57	·66	·31	·53
2. Arith.	·61	·15	·40	·74	·06	·55	·60	·16	·39
3. Sim.	·59	·16	·37	·73	·24	·59	·75	·14	·58
4. Voc.	·70	·28	·57	·88	·03	·78	·76	·16	·60
5. D. S.	·57	·14	·34	·54	·02	·29	·45	·10	·21
6. P. C.	·49	·07	·25	·52	—·25	·33	·49	—·44	·44
7. P. A.	·65	—·10	·43	·64	—·12	·42	·54	—·21	·34
8. B. D.	·58	—·34	·46	·66	—·47	·66	·63	—·41	·57
9. O. A.	·61	—·50	·62	·49	—·49	·48	·52	—·65	·69
10. Cod.	·39	—·10	·16	·45	·00	·21	·68	—·05	·47
11. Verb.	·82	·22	·73	·93	·14	·89	·88	·25	·83
12. Perf.	·79	—·23	·67	·78	—·31	·70	·71	—·41	·67
% Variance	39·2	5·5	44·7	47·9	6·0	53·9	42·3	10·3	52·6

The factors for the 7½ and 13½-year groups were now rotated so that the first factor passed through the Vocabulary sub-test, the second factor being kept orthogonal to it. The factors for the 10½-year group did not require rotation as Vocabulary had a negligible loading on the second factor. All the signs of this factor, however, were reversed to make it correspond with the second factor for the other groups. The rotated factors appear in Table 2.

TABLE 2

ROTATED FACTOR LOADINGS.

	I	II	I	II	I	II
	(7½ years)		(10½ years)		(13½ years)	
1. Comp.	·59	·10	·76	—·06	·71	—·16
2. Arith.	·62	·09	·74	—·06	·62	—·03
3. Sim.	·61	·07	·73	—·24	·76	·02
4. Voc.	·76	·00	·88	—·03	·78	·00
5. D. S.	·58	·08	·54	—·02	·46	—·01
6. P. C.	·48	·12	·52	·25	·39	·54
7. P. A.	·56	·34	·64	·12	·49	·32
8. B. D.	·42	·53	·66	·47	·53	·54
9. O. A.	·37	·69	·49	·49	·37	·74
10. Cod.	·33	·24	·45	·00	·66	·19
11. Verb.	·85	·11	·93	—·14	·91	—·06
12. Perf.	·64	·51	·78	·31	·61	·55
% Variance	34·3	10·4	47·9	6·0	39·4	13·3

IV.—FACTOR INTERPRETATION.

The first factor in each analysis, for reasons of expediency, was made to pass through Vocabulary. This sub-test gives an indication of a child's knowledge of words and, therefore, provides us with a lead for interpreting the factor. However, it would be unwise to label it simply *verbal* since tests like Object Assembly, Block Design, Coding and the composite Performance score also have high loadings on it. A better label would appear to be ' vg '—*verbal-intellectual.* Having anchored the first factor in this way the second factor (orthogonal to it) is found to have neglibigle loadings on each of the first five tests but has consistently high loadings, for each age group, on Block Design and Object Assembly. In consideration of the nature of these two tests, the factor was labelled ' sp '—*space performance.* It should be noted, too, that the three tests, Picture Completion, Picture Arrangement and Coding, have relatively small and very variable loadings on this factor for the different age groups, so that doubt is thrown on the suggestion that the so-called performance tests of the WISC form a cluster, or define a ' performance ' factor.

V.—FACTOR COMPARISON.

In view of the inadequate state of our present knowledge about the standard errors of factor loadings [3] and the suggestion that these errors, for relatively small samples, may be large,[4] it would be unwise to press possible contrasts between the factor loadings for the different age groups too far. No overall trend with age is noticeable, and though the amount of variance accounted for by the first factor, Table 2, is appreciably greater for the 10½-year group than for the 7½-year group, it decreases again when we come to the 13½-year group. For factor II the reverse is true. The general pattern of the factor loadings, however, is alike in all groups ; zero and near zero loadings occur in corresponding positions in each analysis while the higher loadings on each factor are recorded for the same tests.

In view of the small size of the samples—200 in each case—and the fact that standard errors for the loadings are probably quite large, and bearing in mind that the overall pattern of factor loadings in each analysis is similar, it was decided not to set up regression equations for estimating factor scores for each age group separately as the regression weights would probably prove to be

TABLE 3

REGRESSION WEIGHTS FOR ESTIMATING FACTOR SCORES FOR THE FIRST (I) AND SECOND (II) FACTORS, RESPECTIVELY.

	Subtests									
	Comp.	Arith.	Sim.	Voc.	D. S.	P. C.	P. A.	B. D.	O. A.	Cod.
(1) I	·176	·145	·202	·319	·103	·065	·077	·083	(–·017)	·490
(2) II	·103	–·087	–·175	–·208	(–·050)	·125	·092	·349	·541	(·009)
	Rounded Approximations to the above Weights.									
(3) I	·2	·1	·2	·3	·1	·1	·1	·1	0	·1
(4) II	·1	—·1	—·2	—·2	0	·1	·1	·4	·5	0

unstable. Instead the three correlation matrices and the three sets of factor loadings for the ten sub-tests in question were averaged (using Fisher's z-transformation) and a single set of regression weights was obtained ([5], p.221-229). They are given in Table 3 and all except the three in brackets are significant beyond the 5 per cent. level.

Though orthogonality between the rotated factors was retained, estimates of factors scores, being mere estimates, are often slightly correlated. To check on this, the variances of the factor scores and their covariance ([5], p. 238) were estimated. The variances obtained were 0·86 and 0·61 and the covariance 0·075. Hence, the standard deviations (standard errors) of estimated scores on the first and second factors are $\sqrt{(0{\cdot}86)}=0{\cdot}93$ and $\sqrt{(0{\cdot}61)}=0{\cdot}78$, respectively while the correlation between them is

$$r=0{\cdot}075/\sqrt{(0{\cdot}86\times0{\cdot}61)}=0{\cdot}10$$

which is virtually zero.

VI.—USING FACTOR SCORES.

Suppose a child obtains the scores given in the first row of Table 4 on the WISC sub-tests. These scores are now expressed in standard units by subtracting the common mean of 10 from each and dividing the difference by the common standard deviation of 3. The standard scores appear in row two of the table. To obtain the child's score on factor I each standard score is multiplied by its appropriate regression weight (Table 3) and the results added. To ease the calculations involved in getting factor scores approximate weights, as indicated in rows three and four of Table 3, may be used. In this manner the child's score on the first factor is estimated to be 1·49, the standard error of which we already know to be 0·93. In a similar manner his score on the second factor is estimated to be −1·30 and its standard error is 0·78. Since the expected value (or mean) for each factor score is zero, the child's factor scores when compared with their respective standard errors give standardized scores of 1·60 and −1·67, which correspond to the 95th and the 5th percentiles, respectively. We conclude, therefore, that whereas the child's 'vg' factor is high his 'sp' factor is low. If we wish to test whether the child's factor scores differ significantly from each other, we find the ratio of the difference $1{\cdot}49-(-1{\cdot}30)=2{\cdot}79$ to its standard error $\sqrt{(0{\cdot}86+0{\cdot}61-2\times0{\cdot}93\times0{\cdot}78\times0{\cdot}10)}=1{\cdot}15$. This ratio is 2.42 and is significant beyond the 2 per cent. level.

TABLE 4

DATA USED IN FACTOR SCORES EXAMPLE.

	Comp.	Arith.	Sim.	Voc.	D. S.	P.C.	P.A.	B. D.	O. A.	Cod.
Scaled Score ..	10	14	15	17	12	4	16	12	7	11
Standard Score ..	0	1·33	1·67	2·33	·067	−2·00	2·00	0·67	−1·00	0·33

Now, if instead of factor scores we had obtained composite Verbal and Performance scores for the child in question and had then transformed these into Verbal and Performance I.Qs., we would have obtained the values 123 and 100 respectively. From these results we would have concluded that the child's verbal ability is relatively high and his performance ability is average. But the latter result conflicts with our estimate of the child's score on the ' sp ' factor which placed him at the 5th percentile on this factor. This discrepancy is due to the fact that the child's score on the Object Assembly test which is the most important test for defining the ' sp ' factor (see Table 3) is very low, whereas he gets a high score on the Picture Completion test which contributes little to this factor. When situations such as this one are considered and we recall that the composite Performance score is composed of contributions from a very mixed bag of tests, and that this composite score is not independent of the composite Verbal score, the advantages of employing the two virtually uncorrelated factor scores becomes clear. Moreover, the factor scores refer to two relatively ' pure ' and easily appraised aspects of cognitive ability and so greatly facilitate our description of the child's performance.

VII.—REFERENCES.

[1] LAWLEY, D. N. (1951). The maximum likelihood method of estimating factor loadings. *The Factorial Analysis of Human Ability*, Godfrey H. Thomson, 5th ed. London : University of London Press.

[2] MAXWELL, A. E. (1956). Factor models. *J. Educ. Psychol.*, *47*, 129-131.

[3] MAXWELL, A. E. (1957). *A Comparative Study of Some Factor Analytic Techniques.* Ph.D. thesis (unpublished), Univ. of Edinburgh.

[4] MCNEMAR, Q. (1941). On the sampling errors of factor loadings. *Psychometrika*, *6*, 141-152.

[5] THOMSON, G. H. (1951). *The Factorial Analysis of Human Ability*, 5th ed. London : Univ. of London Press.

[6] WECHSLER, D. (1949). *Wechsler Intelligence Scale for Children* (Manual). New York : The Psychological Corporation.

A FACTORIAL ANALYSIS OF THE EVALUATION OF SCHOLASTIC COMPOSITIONS IN THE MOTHER TONGUE*

BY C. REMONDINO

(*Industrial Psychology Laboratory, Fiat School, Turin.*)

SUMMARY. A survey was made of the distinguishable qualities of pupils' compositions, and seventeen qualities were assessed in 230 compositions, both by teachers and non-teachers. A very similar structure of qualities was observed in all sets of marks. On applying factor analysis, it was found that the qualities could be reduced to four main dimensions: Graphic representation, Language usage, Content and arrangement, and Personal aspects of the content.

I.—INTRODUCTION.

THE problem of accurate assessment of scholastic attainment has received useful contributions from general and experimental psychology. Briefly, the main findings concern the personal equation of the examiner, halo effect, and the unsatisfactory validity and consistency of marking.[1] It follows that assessments should, whenever possible, be obtained by means of objective tests or rating scales.[2] These conclusions apply to almost all scholastic situations which use subjective marking. However, the rating of compositions in the mother tongue provides the greatest difficulties. By the term ' composition ' is meant a written exercise which shows the organisation and formation of a series of thoughts about a given topic or theme. This definition covers all the usual school compositions, but excludes those used to test knowledge of history, geography, etc.

The assessment of compositions does not allow, because of its very nature, any recourse to objective methods. In this area the contribution of applied psychology amounts to merely a few general suggestions such as the adoption of an analytical method of marking, reducing the halo effect to a minimum, combining independent markers, etc.

Although improved by the above suggestions, the traditional marking method still holds its own and seems to have no substitute. The marker reads the compositions, corrects grammatical errors, notes down his comments on syntax, style, arrangement and exposition of thoughts, qualifies with some verbal attribute the more complex aspects of the composition and, finally, sums up his conclusion in the form of a quantitative overall judgment. In this case, the material that has to be rated is of an extremely complex nature, containing so many different elements such as grammatical form, content, sentence construction and the subjective contribution by the writer's personality. It is clear that if the above qualities could be separated, we would have achieved something of theoretical interest and practical significance. As a first result, we would be able to mark essays by more methodical and efficient methods.

As far as we know, the most important experiment on these lines is one by Bonnardel entitled "Application de la méthode d'analyse factorielle de Thurstone à l'étude de la notation des copies d'examens,"[3] even if it deals with a type of composition that does not fall within the definition we have postulated.

---

* Paper given at a symposium on educational measurement at the International Congress of Applied Psychology, Rome, in April, 1958, and published in full in *Le Travail Humain*, 1959, 22, 27-40. Translation revised by Dr. W. H. King.

Bonnardel used factorial methods and his results have provided sound theoretical support for the present research.

In the present analysis, we tried to find an answer to the following question : Are there any consistent factors that can be found and how can markers discriminate between these parameters in a normal school composition in the mother tongue ?

The formulation of the question is in itself quite clear but, none the less, we would like to underline the fact that our purpose is different from the aims of the conventional practical experiment in that we shall not consider the validity or consistency of the rater's judgments. Instead, our survey is aimed at pin-pointing the structure of the ratings ; we are trying to assess the qualitative aspect of 'Composition rating' (i.e., the structural arrangement of the parameters for which a rating is normally required) rather than the quantitative accuracy of the ratings.

II.—METHOD OF RESEARCH.

In an attempt to find a satisfactory answer, the experiment was based on the following plan :

(1) A survey among teachers of literary subjects, who collected material for an inventory which was as comprehensive as possible and covered the attributes that raters normally use when correcting school compositions. This inventory was condensed to a limited number of categories.

(2) Rating of a number of compositions by a group of teachers and non-teachers based on the categories found.

(3) Factorial analysis of results.

(4) A cross-check survey with the raters.

The Survey with Teachers.—Approximately twenty teachers of literary subjects were interviewed in different schools (secondary, technical, vocational and lyceums). They were asked : What are the qualities of a school composition that you generally succeed in discriminating ? About a dozen replies were obtained from each teacher and these were grouped, for the sake of simplicity, into a small number of categories without distorting the meaning of the given judgments. These categories were checked independently to ascertain whether they were comprehensive enough to include all the attributes which teachers can detect in their everyday marking. In other words, it was shown practically that a number of judgment-attributes actually used by markers, could be classified in pre-established categories. Only after many cross-checks, all positive, was it decided that the list was satisfactory. Further proof was obtained by comparing this list with a theoretical classification of categories outlined in a modern study,[4] and this list was found to be part of our list which, in its final form, is given in Table 1. The categories are arranged in accordance with a logical sequence suggested by the teachers themselves.

Each category consists of a term followed by a brief definition. All the teachers who have seen this list agree that it is unambiguous and feasible and, consequently, of use in this experiment.

The Rating.—The teachers who were interviewed agreed that seventeen categories could generally be discriminated. A number of markers were then asked to define the 'intensity' of each category in a given essay.

The list was explained to the markers and examples were given of the definitions and these were discussed one by one until it was felt that the meaning was clear and that the categories were accepted and understood by the markers. They were then told how to express their judgment of intensities on a three-point scale. Raters were asked to mark with a + sign the essay in which the

TABLE 1

1. Readability : quality of handwriting—clear, easily read.
2. Aesthetic arrangement : harmonious and pleasing fluency of the letters, good taste in the arrangements.
3. Appearance : clean, carefully presented work.
4. Spelling : measured by number of spelling errors.
5. Morphographic exactness : measured by number of errors in word formation and use of inflexions.
6. Syntax : measured by number of grammatical errors.
7. Organisation of ideas : ordering of ideas and balance.
8. Wealth of ideas : number of ideas, motives, etc.
9. Adequacy of thoughts : quality by which the ideas expressed are pertinent to and adequate for the theme.
10. Factualness : Objective accuracy of the thoughts and facts stated.
11. Comprehensiveness : complete expression of what was intended.
12. Conciseness : use of minimum of words ; without repetitions and redundancies.
13. Language usage : correct use of words.
14. Style : ease and accuracy of sentence construction.
15. Originality : quality by which some part of the writer's personality becomes apparent in the script.
16. Maturity : power of judgment ; ability to reason and to be critical.
17. Imagination : the creative and projective abilities revealed by the script.

parameter considered showed a degree of intensity well above the average ; with a — sign the contrary case and with a O sign the essay in which a distinction from average was not possible.

The markers were warned to be very discriminative and were asked to aim at equal frequencies of +'s, —'s and O's in marking a batch of essays.

The Markers.—Of the ten markers, four were teachers with experience in assessing essays but with different training and background. The other six markers were not school teachers but had average educational ability and had experience in technical or administrative fields.

The Essays.—230 essays were taken on the theme, 'Write about Yourself.' In addition to being one of the most typical kinds of school essays falling within the scope of the definition, it appeared to be capable of arousing, more than many other subjects, a complete range of ideas, personal feelings and subjective attitudes. The essays were written by boys in the 15-16 year age group who were attending courses in a vocational institute (fourth year post-elementary school).

III.—STATISTICAL ANALYSIS.

In the preliminary stage of the work, the intercorrelations of the parameters were calculated using the data provided by all the markers on about one-third of the essays chosen at random. Most of the correlations were computed from a three-fold table, but the tetrachoric technique was used in the remaining cases. The average standard error was approximately 0·13.

The correlation matrix of the markers was analysed using the method of Holzinger and Harman.[5] The average B — coefficient of the identified clusters was approximately 2·50. The results of this analysis are given in Table 2. From this table, still proceeding qualitatively, it was not difficult to derive a structure of clusters from which the markers showed least variation. The following cluster pattern was thus obtained :

(1, 2, 3)
(4, 5, 6)
((7, 11), (8, 9, 10)) 12

$$\left.\begin{matrix}(13, 14)\\(15, 17)\end{matrix}\right\}16$$

TABLE 2

Rater	Clusters
A (teacher)	1, 2, 3 4, 5, 6 7, 8, 9 ....11..........15, 16, 17 10.........13, 14 12
B (teacher)	1, 2, 3 4, 5, 6 7.........11, 12 8, 9, 10 13, 14, 15, 16, 17
C (teacher)	1, 2, 3 4, 5, 6 ....13, 14....16 7...9, 10, 11 8.......................15....17 12
D (teacher)	1, 2, 3 4, 5, 6 7...9, 10, 11..............16 8.....................15....17 12 13, 14
E (not teacher)	1, 2, 3 4, 5, 6.......10 13, 14 7, 8, 9....11, 12........15, 16, 17
F (not teacher)	1, 2, 3 4, 5, 6..............12, 13, 14 7..........................16 8, 9, 10, 11 15....17
G (not teacher)	1, 2, 3 4, 5, 6 7...9, 10, 11..............16 8.....................15....17 12 13, 14
H (not teacher)	1, 2, 3 4, 5, 6 ....13, 14....16 7, 8, 9, 10, 11 12 15....17
I (not teacher)	1, 2, 3 4, 5, 6..................13, 14 7, 8, 9, 10, 11 12 15, 16, 17
L (not teacher)	1, 2, 3 4, 5, 6 7, 8, 9, 10, 11, 12...........16 13, 14 15....17

The parameters included in clusters (1, 2, 3), (4, 5, 6), (13, 14) and (15, 17) follow the above pattern among all the markers.

Clusters (7, 11) and (8, 9, 10) were recognised by nearly half the markers, but became a single cluster with the remainder. Nearly half of the markers assigned parameter 16 to cluster (13, 14) while the remaining markers placed it in cluster (15, 17). With nearly all the markers, parameter 12 was isolated but some markers linked it with cluster (7, 11). The psychological significance of this particular pattern will be discussed in the section dealing with the interpretation of the results.

At this point, it is essential to note :

(*a*) The general validity shown by the cluster pattern in spite of variety of background and experience of the markers.

(*b*) The absence of systematic differences between teachers and non-teacher ratings.

With the help of the information obtained from this preliminary survey, it was possible to apply suitable methods of factorial analysis, using only the marks of the first three teachers (A, B and C). Marker D was not always in close agreement with the first three and his ratings are not included in the calculations that follow.

Ratings were obtained from markers A, B and C on 232 essays and the standard error of zero inter-correlations was 0·08. From these three matrices, a single average matrix was obtained which was analysed using Thurstone's centroid method. Four factors were extracted. McNemar's criterion was used to decide when to stop factorising. A simple oblique-axis structure was then sought by trial and error, using the two different methods of Horst[6] and Boss[7]. Both methods gave very similar results. The cluster pattern and the diagrams obtained from the centroid matrix were both taken into account in fixing the axes (see Appendix, Tables 3-7).

The factorial structure so obtained is indeed very simple and is very close to the ideal case of Holzinger's mono-factor. It should also be noted that the factorial result obtained agrees with the cluster pattern found in the first stage of the investigation and this agreement was most useful in interpreting the final results.

IV.—INTERPRETATION.

Meaning of the factors :

A.—The first factor was heavily loaded with the following parameters :

(1) Readability.
(2) Aesthetic arrangement.
(3) Appearance.

These three parameters cover the exterior form of the composition and so the term GRAPHIC REPRESENTATION may be applied to this group.

B.—The important parameters of the second factor are :

(4) Spelling.
(5) Word formatisn and inflexions.
(6) Grammatical usage (Syntax).
(13) Correct use of words.
(14) Style.

Evidently, each of these five parameters reflects different levels of the same quality displayed by the student and may be classed under the heading LANGUAGE USAGE.

C.—The third factor covers the following parameters :

(7) Organisation of ideas.
(8) Richness or wealth of ideas.
(9) Adequacy of thoughts to subject.
(10) Objective accuracy of facts.
(11) Completeness or comprehensiveness.

These are all different aspects but deal with the CONTENT AND ARRANGEMENT OF THE ESSAY.

D.—The fourth factor included the following parameters :

(15) Originality.
(16) Maturity.
(17) Imagination.

In this group are the more subtle traits that are difficult to define and recognise in essays.

As was clearly in evidence during the cross check, the teachers used these terms to refer to certain aspects of a composition which were very closely linked to the students' personality. According to the teachers interviewed, it would be at least theoretically possible to find two distinct compositions that have the same content, are expressed in similar linguistic and grammatical form with the same quality of handwriting and which, nevertheless, are clearly different. Under these conditions, the distinctive element that may differentiate the two compositions, according to the teachers interviewed, is a ' form ' or ' attribute ' of an absolutely ' personal ' nature.

Originality, maturity and imagination may hence be defined as PERSONAL ASPECTS OF THE CONTENT.

V.—DISCUSSION.

If an objective validatory criterion were available, it would be of great help in discussing the above interpretation of the results ; unfortunately, there is no such criterion and the appraisal of the parameters considered can only be done subjectively except perhaps for a few of the parameters of the second factor.

The idea of proving the validity of the factors found must, therefore, be abandoned. On the other hand, validation is not a strict requirement from the standpoint of what we have tried to measure. What can be done, at this point, is to outline the most interesting conclusions that may be inferred from the results obtained.

(1) The seventeen categories considered are linked by a correlational structure that may be solved factorially to give a simple pattern. Some of these categories have shown very similar factorial content. Although the latter result is no surprise, because it is related to the halo effect whose existence is revealed by the large number of high correlations present in the matrix analysed, it is, nevertheless, interesting to note how even certain parameters which apparently seem to be independent of each other and very easily distinguishable (1 and 2, 1 and 3, 2 and 3 or more so 4 and 5, 4 and 6, 4 and 7, etc.) are instead factorially identical, as shown by the factual evidence. The nature of this experiment does not enable us to say whether the above phenomenon depends upon the subject matter examined or on the subjective impressions of the markers. Bonnardel, in his own research, reached the same conclusion.

(2) A comparison of the ratings of teachers and those who were not teachers is important. In the cases examined it was not possible to trace

structural differences in the ratings assigned by the two types of markers. Hence, the conclusion that may be drawn is that the conceptualization manifested at the time of the marking by the members of the two categories is very much alike, and that, at least among the teachers we have considered, there is no obvious evidence of any professional ' know how ' displayed by the examiners which might be associated with their particular background and experience as teachers. This structural uniformity of judgment appears to be all the more interesting if one considers that, generally speaking, experimental investigations have always shown noticeable deviations even in groups of ratings given by markers with the same training and education. This suggests the following interpretation : while the structure of the ratings is practically unchanged with all the markers, differences are shown when the marker has to quantify the intensity of the attributes identified. This, however, does not exclude the possibility that other differences might exist between teachers and non-teachers ; there may be differences in rate of marking, constancy of results, reliability of marking, etc. Though the mark-remark correlation is a useful method of establishing reliability of markers, we cannot generalise from it because of the rather small number of cases. However, in view of the high level of uniformity found in the separate cluster patterns, and because the ability to discriminate, whether a specific or complex aptitude, ought to be an asset of all minds even if possessed in different degrees, there is reasonable grounds for the acceptance of general validity.

Educational Considerations.—The basic question of this investigation has already been answered, but it may be useful to consider some of the implications of the research from the point of view of the schools.

(1) The results obtained suggest that the marking of essays in the mother-tongue may be simplified by concentration on our four independent factors. These four factors cover most of the elements assessed by the examiners.

(2) From an educational point of view, not all of these factors carry the same weight ; for instance, the one termed Graphic Representation might not be significant at all, especially at certain scholastic levels.

(3) If the marking of a composition is to be a measure of scholastic achievement, it should emphasise the good qualities of the composition that have been taught in school. These qualities may be associated with penmanship, language usage, content and arrangement of ideas. The fourth quality (or factor) which covered originality, maturity and imagination, can hardly be taught, though it can be discriminated.

VI.—BIBLIOGRAPHY.

[1] PIERON, H. (1951). *Méthodologie Psychotechnique*, pp. 124-127, P.U.F.

[2] TITONE, —. (1956). *I problemi della didattica*, p. 363. Pontificio Ateneo Salesiano.

[3] BONNARDEL, R. (1946). Application de la méthode d'analyse factorielle de Thurstone à l'étude de la notation des copies d'examens. *Le Travail Humain*, 8, 130-149.

[4] AGOSTI, —., *et al* (1957). L'espressione. L'insegnamento linguistico. *La Scuola*, p. 168, Brescia.

[5] HOLZINGER, K. J., and HARMAN, H. H. (1951). *Factor Analysis*, pp. 23, 24, Chicago Univ. Press.

[6] HORST, P. (1956). A simple method for rotating a centroid factor matrix to a simple structure hypothesis. *J. Exper. Educ.*, 4, 251-258.

[7] BOSS, J. P. (1957). La recherche directe des facteurs primaires en analyse factorielle. *Bulletin du Centre d'Etudes et Recherches Psychotechniques*, 6, 143-170.

VII.—APPENDIX.

TABLE 3

MATRIX OF INTERCORRELATIONS AND RESIDUALS.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	—	·80	·78	·21	·18	·27	·17	·17	·17	·18	·17	·24	·41	·31	·20	·27	·18
2	·005	—	·86	·26	·23	·15	·13	·08	·12	·17	·22	·28	·34	·31	·19	·25	·18
3	·010	·022	—	·20	·25	·15	·17	·09	·18	·18	·19	·22	·31	·34	·18	·23	·14
4	·042	—·036	·027	—	·70	·61	·31	·27	·27	·28	·27	·25	·52	·49	·26	·32	·29
5	·073	—·013	—·024	·096	—	·74	·34	·26	·33	·35	·32	·33	·57	·58	·29	·34	·30
6	—·051	·021	·035	—·029	·029	—	·49	·37	·40	·39	·38	·36	·71	·73	·39	·43	·37
7	·004	—·002	·021	·010	·026	—·044	—	·56	·65	·56	·54	·37	·53	·52	·46	·58	·46
8	·080	·023	·025	·000	·052	·036	—·031	—	·74	·61	·48	·26	·48	·52	·69	·65	·64
9	·000	—·012	·023	·003	—·011	·012	—·001	·053	—	·63	·64	·46	·54	·55	·53	·58	·54
10	—·019	·001	—·006	·010	—·017	·022	—·020	·008	—·052	—	·62	·43	·52	·57	·46	·53	·46
11	—·077	—·016	—·045	—·031	—·048	—·035	—·011	—·123	—·009	·045	—	·57	·51	·53	·54	·59	·49
12	—·061	—·020	—·077	·006	—·045	—·034	—·032	—·151	·004	·011	·141	—	·48	·45	·35	·41	·29
13	—·032	·013	·044	—·076	—·091	—·011	·005	·031	—·001	·003	—·017	—·043	—	·87	·49	·56	·51
14	·034	·010	—·026	—·106	—·079	·005	·027	·025	—·003	—·041	—·015	—·009	·099	—	·59	·60	·60
15	—·023	—·004	—·035	—·011	—·006	·017	·018	—·092	—·014	·005	·029	·045	—·017	·005	—	·73	·84
16	—·021	—·004	—·005	·006	—·007	·006	—·037	—·026	·024	·014	·015	·029	—·003	—·013	—·003	—	·69
17	—·024	—·022	—·019	—·008	—·029	—·035	·014	—·053	—·034	—·002	·058	·094	—·009	·004	·043	—·019	—

TABLE 4

Matrix of Centroid Factors.

	I	II	III	IV	h²
1	·494	—·479	·462	·217	·734
2	·486	—·511	·576	·205	·871
3	·477	—·504	·508	·268	·811
4	·556	—·332	—·320	—·157	·546
5	·614	—·343	—·387	—·161	·670
6	·688	—·257	—·449	—·169	·770
7	·671	·251	—·168	·153	·565
8	·682	·422	—·118	·047	·659
9	·723	·387	—·140	·291	·777
10	·678	·271	—·125	·238	·605
11	·690	·315	·062	·127	·595
12	·566	·065	·062	·107	·340
13	·826	—·175	—·214	—·070	·764
14	·845	—·110	—·196	—·165	·792
15	·709	·370	·191	—·388	·827
16	·760	·299	·110	—·159	·704
17	·700	·339	·124	—·392	·774

TABLE 5

Final Matrix, After Rotation.

	A	B	C	D
1	·72	—·10	—·09	—·17
2	·80	—·16	—·15	—·15
3	·79	—·16	—·08	—·22
4	—·07	·59	—·04	·06
5	—·11	·66	—·01	·06
6	—·18	·70	·07	·10
7	—·06	·13	·49	·07
8	—·17	·09	·50	·24
9	—·03	—·01	·68	·02
10	·00	·04	·56	·01
11	·05	—·03	·45	·17
12	·14	·04	·26	·07
13	·01	·49	·16	·13
14	—·06	·52	·13	·24
15	—·18	·16	·07	·67
16	—·06	·13	·23	·44
17	—·21	·22	·06	·65

TABLE 6

Transformation Matrix.

	A	B	C	D
I	·125	·295	·293	·214
II	—·482	—·385	·536	·376
III	·665	—·662	—·297	·212
IV	·556	—·571	·734	—·876

TABLE 7

Matrix of the Cosines of the Angles between Primary Vectors.

	A	B	C	D
A	·999			
B	—·054	1·000		
C	—·011	—·342	1·000	
D	—·501	·278	—·442	·999

AN HISTORICAL AND CRITICAL ACCOUNT OF THE ACCOMPLISHMENT QUOTIENT IDEA

By A. R. CRANE

(*Teachers' College and University of New England, Armidale, New South Wales, Australia.*)

It is probable that the A.Q. was talked about in university lecture rooms before 1920 when it first appeared in the American literature. It was in that year used in two unconnected investigations, one by Monroe and Buckingham and the other by R. H. Franzen. At the outset it will, therefore, be worthwhile to give a critical examination of these people's ideas since only in this way can we fully understand the inverted pyramid of confusion that has been built upon them.

Monroe and Buckingham produced the Illinois Examination, a standardised group test battery which included general intelligence, arithmetic and silent reading. The norms for the reading and arithmetic tests were derived from the performances of pupils of the same mental age as indicated by the general intelligence component of the battery. These norms they expressed in Achievement Ages from which each pupils' Achievement Quotient (A.Q.) was calculated by dividing by his Mental Age. The collaborators argued in this way: "A pupil's mental age represents his capacity to learn. If his Achievement Age is greater than his Mental Age he has achieved more than the average pupil of his Mental Age under average school conditions . . . If a pupil's Achievement Age is less than his Mental Age we have evidence that he has not achieved as much as he should" ([21], p. 16).

Since this was the prototype of all such statements, it deserves careful attention. Notice particularly three things:

(1) The postulate that a pupil's Mental Age represents his capacity to learn school tasks. Although this idea is co-eval with the intelligence testing movement, it has seldom been pointed out that it involves many difficulties. For example, Mental Age is a statistic and not a measure. It is not even a score. It represents the *average* performance of a number of children of a certain Chronological Age. There is, therefore, nothing precise or definite about it. This confusion about the nature of the Mental Age (or of norms of any sort for that matter) has always bedevilled the whole mental testing movement, where there is still a widely accepted idea that there is some sort of causal relationship between a child's score on an intelligence test and the level of his performance in school. How often do we hear it said that a child is "backward *because* his I.Q. is 80" or that "he should be doing better *because* his I.Q. is 130." It would be just as valid to say that a pupil's I.Q. is 80 *because* he is backward.

(2) Notice secondly the statement that "if Achievement Age is greater than Mental Age (i.e., Achievement Quotient greater than 100) it means that the pupil has achieved more than the average pupils of his Mental Age under average conditions." This is quite an unimpeachable statement, but it is also quite inconsistent with statement (1) above. If Mental Age represents a child's capacity to learn (as statement (1) says it does) then if Achievement Age is greater than Mental Age it means that he has achieved more than he is capable of achieving. This paradox has been a very worrying problem over the years.

Those who treat Mental Age as a measure of capacity are obviously puzzled and upset by such cases and dedicate their efforts to explaining them away. Those who treat Mental Age as a statistic accept such cases as a matter of course since clearly a pupil can perform better than the average pupil of his Mental Age.

(3) In the third place, notice the interpretation of cases where Achievement Age is less than Mental Age. " We have evidence that he has not achieved as much as he *should*." This follows directly from statement (1) but is quite inconsistent with statement (2). If they had said that a case of Achievement Age less than Mental Age meant that the pupil had achieved less than the average pupil of his Mental age no exception could be taken and the future of the A.Q. idea might have been less disastrous.

As it is, there can be no doubt that Monroe and Buckingham's A.Q. was born in confusion. This was not improved by Franzen, the other pioneer, who independently developed an A.Q. (Accomplishment Quotient this time) which was somewhat different from the Achievement Quotient, but which had the same idea behind it. Franzen used Educational Ages in reading and arithmetic derived from norms based on *chronological age* groups. The E.A./C.A. ratio he called the Educational Quotient and argued : " Since the I.Q. is the potential rate of progress, the ratio E.Q./I.Q. gives the percentage of what that child could do to what he has actually done" ([12], p. 435). (Although these are Franzen's actual words his meaning would have been clearer if he had said " the ratio E.Q./I.Q. gives the percentage of what that child has actually done to what he could do ").

He stated that under ideal educational conditions the correlation between E.Q. and I.Q. would become unity and the A.Q. would then be 100 for all pupils. " It is, therefore, safe," he wrote, " to assume that the optimum A.Q. is 100." But what of cases where A.Q. was greater than 100 ? Franzen fell into the trap which is opened as soon as M.A. or I.Q. are interpreted as measures of potential achievement. Where such cases occurred, Franzen argued, it was evidence that the achievement test norms were too low—they were not what he termed 'pushed norms,' that is, norms based on the performance of pupils achieving at their maximum level.

A variation in the usage of A.Q. was introduced in 1921 by Madsen,[20] who was an administrator interested in the relative efficiency of teachers, not a psychologist concerned with the individual performances of children. He used A.Q. to describe the *average* performance in a school subject of a class group of pupils relative to the *average* I.Q. of the class. It was in this way that the A.Q. was in the future to be commonly used by American administrators.

In 1922, Franzen expanded his original article and made a radical change in his terminology. His former E.A. he now called Subject Age and the E.Q. became the Subject Quotient. The ratio S.Q./I.Q. became the 'Subject Ratio ("previously called the accomplishment quotient" ([13], p. 5), he laconically writes). Then he introduced an Accomplishment Ratio which was an average of the Subject Ratios. He still maintained that the S.Rs. should be 100 before a pupil's performance in any single subject could be considered satisfactory. Not only did this add confusion to terminology, but it introduced another confusion all its own—that of whether the A.Q. idea could be applied to single subjects or only to the total school performance. It was in this second sense that Burt was independently of Franzen to use the idea in England.

Franzen noticed that children of lower I.Q. tended to have Subject Ratios higher than their I.Qs., but he was still satisfied that the explanation lay in the achievement test norms. He did not recognise it as, in part at least, the inevitable result of regression between his two imperfectly correlated variables, Mental Age and Subject Age.

The A.Q. idea which developed from the sources mentioned, was immediately given an enthusiastic reception in educational circles. McCall, for instance, in his influential textbook *How to Measure in Education*, first published in 1922, wrote: "The A.Q. is the most exact present-day measure of the efficiency of study, instruction and supervision. It is the best index of what pupils need special attention and spurring, of what pupils need restraining perhaps and of what pupils need to be let alone" (19, p. 86). "It is," he enthusiastically claimed, "a device for converting pupils, no matter what their intelligence, into standard pupils." (19, p. 152).

The idea now quickly gathered momentum. Pintner discovered that not only were pupils of below average intelligence doing better than expected in attainments tests, as Franzen had reported, but also conversely, the bright pupils were as a group not achieving 'up to expectations.' Here we have the regression phenomenon in full flower. (Franzen had seen but half of it). Pintner's data could not be explained away on the basis of Franzen's argument that the attainments test norms were too low. They could hardly be too low for the below average group and at the same time be too high for the above average pupils. Pintner drew his own conclusions which were to be accepted, echoed and re-echoed through most of the subsequent literature. "It is evident, therefore," he wrote, "that the tendency of the school is to push ahead the mentally slow and at the same time to neglect the bright as soon as they have achieved average work." (26, p. 166).

In 1926, Peters made an attempt to apply the A.Q. technique to work in high school and college and his data 'confirmed' the then well accepted idea that the bright were being neglected and he maintained that this was happening even at the secondary and tertiary level. "There is a widespread conviction," he wrote, "that unlike the lower schools, colleges cater to their more brilliant students at the expense of their duller ones . . . The writer shared this conviction and fully expected, as a consequence, a high positive correlation between brightness and accomplishment. But the opposite turned out to be the case . . . apparently the parable of the talents, when told to college students needs to be reversed." (24, pp. 107-8). Strangely enough, it was not until 1939 that it was pointed out that statements like this flew in the face of common sense. "Observation," wrote Tiegs, "does not support the claim that dull pupils work more faithfully or more diligently as a group than do the bright; in fact, there is evidence to believe that the opposite is true." (33, p. 332).

Soon after the article by Pintner (which has been referred to above) appeared in 1922, Torgerson[34] presented much the same kind of data, but in considering possible explanations he mentioned something new, the effect of grade placement upon a pupil's performance in attainment tests. It was not until three years later that Florence Goodenough, (14, p. 298), made this point amply clear. "Learning," she pointed out "is a function of the time one has been learning as well as of ability." By this time Torgerson, as the result of further data he had collected, was convinced that "the reason the accomplishment quotient bears an inverse ratio to intelligence quotients is not because the relationship is an inherent one, but rather that it has been produced through the artificial system of the traditional school with its improper grading and classification" (35, p. 133-4).

During this time, Burt, in England, was producing his now classical *Memoranda on the Relations Between Mental Ability and Educational Attainment*. In these reports he did not use A.Q. as such, but he was much concerned with the discrepancies he found between his Mental Ratio and Educational Ratio, which he based on a pupil's actual grade placement relative to the grade considered 'proper' for a pupil of his chronological age (3, p. 15).

Amongst pupils of M.R. 85–100 Burt noticed "a discernible effort, an effort by no means sterile, to coax and coach these milder dullards to a grade more fitted to their actual age." Then followed a statement which was to colour all future English thought on this matter—" The mental legacy which they inherit is slender, yet so judiciously have their teachers invested and improved it that the accumulated interest now exceeds the capital" ([4], p. 176). He also noted that " the slightly above average children, despite superior talent, are largely kept back scholastically" ([5], p. 178), and he claimed that his data " unmask a strong disposition to level a child's school work up or down towards a common standard for his age" ([8], p. 178). By a different route Burt had come to much the same conclusions as his American contemporaries Pintner and Torgerson. It is equally important to notice that Burt found that " out of the entire group ($N=689$), six receive an E.R. 30 per cent. or more below their M.R. The reverse it may be noted, never occurs" ([4], p. 178). This shows that Burt was in agreement with those like Franzen who considered an A.Q. above 100 as an abnormal phenomenon requiring a special explanation.

It is interesting to notice that by the mid-1920's English and American workers had quite independently come to much the same conclusions about the relationship between ability and achievement, At about this time in America, there began a series of mildly critical examinations of the A.Q. Chapman,[9] for example, pointed out the high unreliability of this statistic which was based upon two other statistics of often unspecified reliability. However, it was not until 1928 that it was shown that, since ability and achievement were imperfectly correlated, it was a statistical necessity for dull children as a group to have E.Qs. above their I.Qs. and, conversely, for bright children as a group to have E.Qs. below their I.Qs. This was done by Wilston who, in an excellent article, summed the matter up succinctly : " Even by an inordinate effort the bright child cannot raise his A.Q. to unity and even if the very dull student were as lazy as the bright are commonly believed to be, the A.Q. of this group would still be gratifyingly large " ([38], pp. 2-3).

This convincing demonstration of a basic weakness in the A.Q. idea did not halt its momentum. Nygaard[22] for example, admitted the truth of all Wilson said and then proceeded to devise a " corrected A.Q." using a formula which he claimed cancelled out the influence of regression.

The situation rested in this state on both sides of the Atlantic until 1936, when Schonell[29] contributed a section on " The Education of the Backward Child " to the *British Yearbook of Education*. He mentioned the A.Q. in a footnote where a warning was sounded of inaccuracies which were likely to arise first from a neglect of C.A. and second, through insufficient evidence of the relative rates of progress of children of different I.Qs. Schonell's footnote is interesting for two reasons—the absence of any recognition of the regression problem as it affected the A.Q., and the use of the term ' retarded ' to describe children whose E.A. was below their M.A. The A.Q. idea in England gained its greatest currency under this concept of retardation.

The publication, in 1937, of Burt's *The Backward Child*, was a momentous event. It had much to say about the Achievement Ratio, and here I must quote at some length. " Perhaps" Burt writes, " the most relevant comparison of all turns on the relation between the child's school attainments and his inborn intellectual capacity. If a child's A.R. is exactly 100, then teaching is keeping pace with mental growth ; the child, it may be inferred, is learning his lessons up to the full limit of his inborn capacity. Achievement Ratios rising above 100 are occasionally met with. At first sight they may seem something of a paradox and the examiner may begin to question the accuracy of his tests. How, he will ask, can a child possibly show achievements which are greater than his ability ? These high ratios, however, are undoubtedly rare. They occur sporadically in a few young bookworms who show an extra zeal or talent in academic work, but less practical shrewdness and common sense. They occur rather more frequently, in dull youngsters who have been assiduously coached by a good teacher and have been laboriously brought on, till their acquirements in the more mechanical subjects rise above what other pupils of their own meagre powers would ordinarily reach " ([5], p. 35).

Further on, he emphasised this same idea: " Capacity must obviously limit content. It is impossible for a point jug to hold more than a pint of milk, and it is equally impossible for a child's educational attainments to rise higher than his educable capacity " ([5], p. 447).

A similar viewpoint was taken by W. P. Alexander who, three years later, stated that " the dull child will be backward, indeed, from a purely educational point of view they *should* be backward. They should be as backward as they are dull" ([1], p. 45). He emphasised his meaning by printing "should" in italics.

In 1942 Schonell's *Backwardness in the Basic Subjects* was first published. In this book much was made of the difference between backwardness and retardation. " Retarded children," Schonell wrote, " have intellectual powers which are superior to their scholastic achievements . . . They represent cases of improvable scholastic deficiency" ([30], p. 61). In effect, Schonell is in complete agreement with Burt's statement that " under normal conditions the teacher may gauge his own efficiency as well as the child's by the success in keeping the A.R. of all his pupils close to 100 per cent." ([5], p. 36). This is so even though Schonell showed caution with respect to the use of the A.Q.

With the enormous prestige of Burt's statements and the eminent practicability of Schonell's tests in the classroom, the retardation idea was away to a good start in England. It tended to dominate the thinking of teachers, educational psychologists and clinicians. In the training of all these workers much was made of the difference between the child who was ' backward ' because he was dull and the child who was ' retarded ' because he was not ' working up to his ability.' Retardation was looked upon as a remediable condition but backwardness was irremediable. This point of view was well expressed by Schonell in 1935: " . . . we can distinguish . . . three groups of pupils who exhibit some form of scholastic disability:

(*a*) The innately dull who, are backward because of inferior intellectual powers and hence need teaching differing qualitatively and quantitatively from that accorded to normal children.
(*b*) Those of unimpaired general intellectual powers who are backward in school work and who simply need continous schooling, individual assistance and special coaching to overcome this handicap.
(*c*) The supernormal pupils whose disparity between I.Q. and A.Q. should hardly be considered as backwardness, a more suitable term being retardation; their backwardness is more apparent than real. Such scholars require increased private study and a fuller curriculum to extend them to the limit of their intellectual capacities " ([28], p. 476).

The almost inevitable result of this kind of approach was to concentrate attention and interest upon the retarded children with the consequent comparative neglect of the backward.

It is interesting to notice that this progressive concentration on the idea of retardation was happening in England at the time when the allied idea of the A.Q. had all but been discarded in America because of the widespread recognition of its many inherent weaknesses. For example, the *Thirty-Fourth Yearbook of the National Society for the Study of Education*, did not give the A.Q. a single mention in their whole volume on Educational Diagnosis which appeared in 1935. At about the same time American textbooks in educational psychology were either making no reference to the A.Q. at all (e.g., Griffiths[16] and Judd[18],) or were urging extreme caution in its use (e.g., Tiegs[33] and Greene[15]). This last writer could quite accurately report that " the A.Q. has nearly gone out of use " ([15], p. 251).

When the idea of retardation did come under criticism in England, it was not the result of this trans-Atlantic scepticism but the outcome of a train of thought that had been developing within English educational psychology itself. From the late 1940's there was a rising body of opinion which called for a re-thinking of the meaning of intelligence test results and for a re-assessment of the distinction which was being so clearly made between ability and attainment ; ability was inborn and measured by intelligence tests, attainment was acquired and was measured by achievement tests. For example, Vernon[36, 37] wrote, in October, 1956 : " The inescapable conclusion is that this whole conception of retardation relative to intelligence is so unstable, so dependent on the particular tests employed, as to be almost meaningless psychologically and educationally." ([37], p. 1277).

At the same time, quite independently, the National Foundation for Educational Research had administered a non-verbal intelligence test and tests of reading and arithmetic to some 3,000 children. The results were published by Pidgeon and Yates[25] in November, 1956, and showed that the distribution of reading A.Qs. was very nearly normal. This should have caused no surprise since the normality of at least some A.Q. distributions had been demonstrated several times in America. For example, Popenoe[27] as early as 1927, presented data secured by testing 600 pupils with a verbal intelligence test and standardised test of reading, arithmetic and spelling. The A.Qs. calculated from an Educational Age which was the average of Arithmetic, Reading and Spelling Ages gave a near normal distribution. For this reason (amongst others) its use was ' finally discontinued ' in the Los Angeles City Schools. In 1930 Coy,[10] using the Stanford Achievement Test and a group verbal intelligence test with over 400 pupils secured much the same result. A more important study was that by Cureton[11] in 1937 because he used data from both Binet and group verbal Otis tests with the Stanford Achievement Test. The distributions of both Binet and Otis A.Qs. were normal. The standard deviation of Binet A.Qs. was 8·61 and of Otis A.Qs. was 5·39.

It seems more than likely that this American work was not known to the N.F.E.R. investigators since they appeared to be surprised by their own data. " It is surely necessary to question the validity of many of the assumptions that underlie the techniques for the diagnosis of backwardness and retardation that are nowadays employed in classrooms and clinics " ([25], p. 26).

Sir Cyril Burt criticised the Foundation's data on the ground that they were based on a single non-verbal group intelligence test and a single test of reading. He defended his interpretation of the A.Q. by the use of a metaphor. " The A.Q. in fact is rather like the ratio used by engineers to measure the efficiency of an engine or other machine. The efficiency of a steam engine or an electric motor is expressed by the ratio of the amount of work actually done to the amount of energy (capacity to work) supplied to the machine" ([6], p. 913). He then qualified his statement : " However, unlike the efficiency ratio, a pupil's A.R. may, in certain cases exceed 100 per cent. A child with an exceptionally good memory, high verbal capacity and a passion for story books will nearly, always achieve a higher attainment in reading than an average child with the same amount of general ability."

Then began a vigorous correspondence in the pages of *Education*, the official organ of the Association of Education Committees. Burt was supported by Babington Smith[32] who argued that Pidgeon and Yates had " missed the point " because Burt's argument had rested on data from the Binet and not from group tests. He ignored Cureton's work which suggested that A.Qs. calculated from Binet I.Qs. were spread more (not less) widely than A.Qs. based on a group test I.Q.

Peel wrote to say that " most of us know that A.Rs. are distributed more or less symmetrically " ([23], p. 970), and in doing so, contradicted Burt. Many other well-known English psychloogists wrote in support of Pidgeon and Yates, but at the close of the correspondence Burt repeated that he was convinced that ' over-achievers ' are far less common than ' underachievers ' ([7], p. 507).

In March, 1957, the N.F.E.R. gave Burt, Peel and Schonell an opportunity to state their opinions in full. Burt reiterated his former arguments whereas Schonell opened his article by declaring that " the distribution of the ratios of two test scores could not possibly be asymmetrical "[31] That is where, so far as the published material is concerned, the matter rests at the time of writing (October, 1958).

In this same number of the N.F.E.R. *Bulletin*, an article by the Belgian psychologist, Hotyat, had some very wise things to say on the matter. " Belief in the notion that it is impossible for attainment to surpass intelligence is in some measure the price paid for the remarkable progress of educational psychology in England. The method devised by Spearman and his followers was so fruitful and aroused so much enthusiasm that, little by little, the methaphorical hypotheses he used to give concreteness to the notion of ' g ' impregnated this number." ([17], pp. 29-30). In other words, the ' milk and jug ' image had taken control of English thinking about the relationship between ability and attainment.

At this point it should be made clear that this critical survey is not a denial of the existence of children who may be described as educationally retarded—that would be to fly in the face of common sense. What is called into question is the validity of much current practice in the detection of those who are retarded and the massive confusion of thought in which the concept of retardation has become inextricably snared. The Editor of *Education* summed the matter up in May, 1957. " Unfortunately not only are the theoretical assumptions still in dispute, but there seems to be no agreement on the practical implications of the controversy " ([39], p. 787).

One might pardoned for asking whether the A.Q. is one of those compelling errors which have provided a fertile soil in which science has its roots. Chemistry grew out of Alchemy, astronomy from astrology. Will a sound psychology of achievement emerge out of the A.Q. ?

REFERENCES.

[1] ALEXANDER, W. P. (1940). *The Educational Needs of a Democracy*. London : Univ. of London Press.

[2] ALEXANDER, W. P. (1957). Reply to letter from Pidgeon and Yates. *Education*, 109, p. 408.

[3] BURT, C. (1917). *The Distribution and Relations of Educational Abilities*. London : King.

[4] BURT, C. (1921). *Mental and Scholastic Tests*. London : King.

[5] BURT, C. (1937). *The Backward Child*. London : Univ. of London Press.

[6] BURT, C. (1956). Letter to the Editor. *Education*, 108, p. 913.

[7] BURT, C. (1957). Letter to the Editor. *Education*, 109, p. 507.

[8] BURT, C. (1957). The relation between ability and attainment. *Bulletin of the National Foundation for Educational Research in England and Wales*. No. 9, pp. 1-17.

[9] CHAPMAN, J. C. (1923). The unreliability of the difference between intelligence and educational ratings. *Journal of Educ. Psychol.*, 14, pp. 103-106.

[10] COY, G. L. (1930). A study of the various factors which influence the use of the accomplishment quotient as a measure of teaching efficiency. *Journal of Educ. Res.*, 21, pp. 29-42.

[11] CURETON, E. E. (1937). The accomplishment quotient technic. *Journal of Exper. Educ.*, 5, pp. 315-326.

[12] FRANZEN, R. H. (1920). The accomplishment quotient. *Teachers' College Record*, 21, pp. 432-440.

[13] FRANZEN, R. H. (1922). The accomplishment ratio: a treatment of the inherited determinants of disparity in school products. Columbia Teachers' College Contributions in Education, No. 125.

[14] GOODENOUGH, F. (1925). Efficiency in learning and the accomplishment quotient. *Journal of Educ. Res.*, 12, pp. 297-301.

[15] GREENE, E. B. (1941). *Measurements of Human Behaviour.* New York: Odyssey Press.

[16] GRIFFITHS, C. (1938). *Psychology Applied to Teaching and Learning.* New York: Farrar and Reinhart.

[17] HOTYAT, F. (1957). Ability and attainment: a note. *Bulletin of the National Foundation for Educational Research in England and Wales.* No. 9, pp. 28-31.

[18] JUDD, C. H. (1939). *Educational Psychology.* Boston: Houghton-Mifflin.

[19] MCCALL, W. A.: (1922). *How to Measure in Education.* New York: Macmillan.

[20] MADSEN, I. N. (1921). Interpreting achievement in school in terms of intelligence. *American School Board Journal*, November issue.

[21] MONROE, W. S., and BUCKINGHAM, B. R. (1920). *The Illinois Examination.* Bloomington: Public School Publishing Co.

[22] NYGAARD, P. H. (1928). A revised accomplishment quotient. *Journal of Educ. Res.*, 18, p. 87.

[23] PEEL, E. A. (1956). Letter to the editor. *Education*, 108, p. 979.

[24] PETERS, C. C. (1926). A method of computing accomplishment quotients on the high school and college levels. *Journal of Educ. Res.*, 14, pp. 99-111.

[25] PIDGEON, D. A. and YATES, A. (1956). The use of tests in the class-room. The relationship between ability and attainment—an examination of current theory. *Bulletin of the National Foundation for Educational Research in England and Wales*, No. 8.

[26] PINTNER, R. (1922). The significance of intelligence testing in the primary school. *Twenty-first Yearbook of the National Society for the Study of Education.*

[27] POPENOE, H. (1927). A report of certain significant deficiencies of the accomplishment quotient. *Journal of Educ. Res.*, 16, pp. 40-47.

[28] SCHONELL, F. J. (1935). Diagnostic tests for specific disabilities in school subjects *Yearbook of Education*, pp. 475-480. London: Evans.

[29] SCHONELL, F. J. (1936). The education of backward children. *Yearbook of Education*, pp. 790-797. London: Evans.

[30] SCHONELL, F. J. (1942). *Backwardness in the Basic Subjects.* Edinburgh: Oliver and Boyd

[31] SCHONELL, F. J. (1957). The relationship between ability and attainment. *Bulletin of the National Foundation for Educational Research in England and Wales.* No. 9, pp. 18-22.

[32] SMITH, B. B. (1956). Letter to the editor. *Education*, 108, pp. 978-979.

[33] TIEGS, E. W. (1939). *Tests and Measurements in the Improvement of Learning.* Boston: Houghton-Mifflin.

[34] TORGERSON, T. L. (1922). The efficiency quotient as a measure of achievement. *Journal of Educ. Res.*, 6, pp. 25-32.

[35] TORGERSON, T. L., and SHUMAN, I. (1925). The variability of the accomplishment of pupils of the same mental level. *Journal of Educa. Res.*, pp. 132-136.

[36] VERNON, P. E. (1955). The psychology of intelligence and "g." *Bulletin of the British Psychological Society*, No. 26, pp. 1-14.

[37] VERNON, P. E. (1956). Dullness and its causes. *Times Educational Supplement*, Oct. 26th, p. 1277.

[38] WILSON, W. R. (1928). The misleading accomplishment quotient. *Journal of Educ. Res.*, 17, pp. 1-10.

[39] ——— (1957). Editorial: The debate continues. *Education*, 109, p. 787.

(*Manuscript received September 30th*, 1958.)

The Accomplishment Quotient: A Reply.

BY CYRIL BURT.

Mr. Crane, I fancy, has been busily engaged in flogging two dead horses. His main criticism, if I understand him rightly, is directed against those who divide the educationally backward into two mutually exclusive types—the 'inherently dull' and the 'merely backward' (to use the old Board of Education terms), and then suppose that, while 'retardation' is a 'remediable condition,' backwardness due to inherent dullness is 'irremediable'. This notion he attributes to "the prestige of Burt's statements and the practicability of Schonell's tests." However, if he will look again at the book he cites (*The Backward Child*), he will find that it is one long protest against the idea that "there are only two main causes and two distinct and non-overlapping types." "Time after time," I said, "it proves almost impossible to decide whether a child's backwardness springs chiefly from innate weakness, or from environment handicaps, or from both conspiring together: each child must therefore, be considered as a unique individual" (p. 14; cf. p. 564).

The distribution of the 'achievement ratio' is a theoretical problem, rather too technical for full discussion here. In *The Backward Child*, he asserts, I "had much to say about the Achievement Ratio." Actually I said very little about it. It was included and illustrated for the sake of completeness, but never used in the descriptions of individual cases. And I fully agree with Mr. Crane in deprecating its use. Curiously enough he does not mention the chief objection. Statistically ratios are always unsatisfactory measures; and a ratio of two ratios the most unsatisfactory of all. As I myself stated, "to compare a child's actual attainments with his potential ability it is much better to take *differences between standard measures*" (p. 36: cf. *N.F.E.R. Bulletin*, x., 1957, p. 9: how Mr. Crane manages to read this article as a *defence* of the A.Q. passes my comprehension). He states that Pidgeon and Yates "showed that the distribution of reading A.Q's was very nearly normal"; but he overlooks the fact that their distribution is for a *logarithmatic* function of the A.Q's, not for the A.Q's themselves—the transformation being used to correct for the asymmetry. Mr. Crane seems to share the popular notion that, because two variables (e.g., mental age and educational age) are normally distributed, not only their difference, but also their ratio will be normally distributed: this, however, is an error (see Kendall, *Advanced Theory of Statistics*, I, p. 253, and ref.).

As his paper is partly a 'historical account' of the idea of the A.Q., may I correct one minor misunderstanding? After describing the conclusions reached in the U.S.A. by Pintner, Peters, and Torgersen in 1922-6, Mr. Crane says that, in my L.C.C. memoranda on *The Distribution and Relations of Educational Abilities* (which he misquotes as *The Relations between Mental Ability and Educational Attainments*), I was "much concerned with the discrepancies between the Mental Ratio and the Educational Ratio." Now there is no reference to the 'mental ratio' throughout that volume, except on the page he quotes where a footnote *criticizes* the use of the I.Q.; and in any case that page refers solely to defectives in special schools. Moreover, the report was published in 1917—long before the American work on the A.Q., with which he declares (quite wrongly) that I "was in agreement." In a later Report (1921), as he says, I pointed out a tendency in schools to "level the child's work up or down towards a common standard for his age." But this statement was put forward, not as a theoretical conclusion agreeing with the doctrines of the American advocates of the A.Q., but as a record of an actual fact—a point of history which every English teacher whose memory goes back to the "bad old days" will confirm. At that date the official regulations assumed that children would be promoted annually by age; classes were enormous; teaching was only practicable if nearly all the pupils were working at much the same level; and, though 'payment by results' had by then been abolished, inspectors still required every child to conform as closely as possible with 'standards' laid down in the Board of Education code. Possibly Mr. Crane is really referring to school conditions in Australia; if so, I should not for a moment wish my conclusions to be applied to schools in other countries.

The paradoxes in the distribution of A.Q's, on which he lays so much stress, doubtless arise chiefly as a result of special abilities or disabilities, which he wholly disregards. But I am entirely at one with him in deprecating an excessive reliance on the results of mental tests alone. As I emphasized in the preface to the book he cites, "tests can be only the beginning, never the end, of the examination of the child."

RESEARCH NOTE

Note on the Use of New-Type Tests for Improving the Quality of Discussion in Discussion Groups.

By A. J. M. FLOOK

(*Department of Psychology, Queen's College, Dundee, University of St. Andrews*).

So long as spontaneity remains one of the main attractions of a good discussion, the likelihood of occasional misfirings will be a natural hazard in discussion groups. However, premeditated intervention on the part of the discussion leader may improve the general quality of discussions, and one attempt to do so without extinguishing spontaneity by too much forethought and control is reported below.

PROCEDURE.

In a course on Management, four two-hour periods were included as an introduction to industrial psychology. It was decided that the first three periods should take the form of lecture-discussion meetings and that all of the last should be reserved for discussion. A ' true/false ' test was given to the class at the end of the third period, one reason for this being to provide ready-made material for use if spontaneous discussion failed in the fourth period. This particular form of test was chosen in the hope that the debating society flavour of the ' true/false ' item might help to create a suitable atmosphere of excited controversy. The test's twenty-seven items (taken with very little modification from N. R. F. Maier's *Psychology in Industry*, 1946) mostly concerned general issues, partly because of their suitability as discussion material.

The results of the test (which took about 10 minutes) did not provide a sure basis for predicting which of the five broad topics covered by it would probably provide the most profitable subjects for discussion in the last period. A better indication was given by the distribution of answers to individual items, which showed that nine of them had given rise to extensive disagreement, either within the group or between it and the ' official ' answers. Of these nine, five looked promising as starting points for discussion, and it was decided to fall back on them if the group did not spontaneously generate a satisfactory discussion.

A further decision was to exploit the divergence between the opinion of the group and the ' official ' answer to item 3, which ran : " In their assessment of how satisfactory a job is, workers attach most importance to high wages and short hours." Four out of the five members of the class marked this ' true,' and this show of resistance was symptomatic of the group's scepticism, at previous meetings, regarding the industrial psychologist's approach to the problem of incentives in industry. Consequently, particular attention was given to preparing for discussion of this item.

The fourth meeting began with an affably worded inquiry as to the grounds on which the items had been marked as being true or false. The opportunity was taken to explain that, unlike most new-type tests, this one was a test of judgment rather than of knowledge, and that the official answers were to be taken merely as crude indications of the likely consensus of opinion among industrial psychologists. The discussion leader concluded these preliminaries by specifying the five items which seemed likely to yield the most fruitful discussion. For most of the next half-hour, however, discussion centred on two other items, but was then drawn round, without opposition, to item 3. The findings of investigations into factors affecting the choice of jobs were described, and it was emphasised that these studies, despite their shortcomings, constituted a scientific approach to the question of industrial incentives. The discussion thus aroused left the impression that the group, although still not altogether repentant, had developed a greater understanding of the complexity of such issues.

COMMENTS.

Using the ' true/false ' test in this way had, it was felt, had the following results :

(*a*) It had provided a practical demonstration of a ' new-type ' test in action.

(*b*) It had provided a rough measure of how much had been assimilated in the earlier meetings.

(*c*) It had added to the liveliness of the third and fourth meetings.

(*d*) It had drawn attention to the group's reluctance to accept part of the content of the lectures.

In short, use of the test had made a valuable contribution to the course, in particular by improving the quality of discussion. It seems reasonable to suppose that the test had this latter effect by unequivocally committing all members of the group to a definite point of view, which they were, therefore, strongly motivated to defend in discussion. Ordinarily, when discussion time comes, the members of a group may, if they so desire, preserve a stony silence with impunity : they can ' sit on the fence ' without loss of self-esteem or of standing in the eyes of others. On this occasion, the test had toppled them off the fence, leaving them so ' ego-involved ' that it was not open to them to react with the traditional dumb detachment. This effect would have been even greater, presumably, if an answer to every item had been insisted upon.

Several other points of interest emerged :

(*a*) The test results were tabulated so as to show the declared attitude of each individual to every item. One use for such a table is as an aid to the leader in drawing the more reserved members effectively into the discussion. Another is as a check on the discussion being side-tracked by self-assertive individuals. If, as seems likely, most persons prefer to discuss the items they fail to get ' right,' the table identifies those topics which will probably interest the majority and the leader can use these as landmarks in steering the discussion.

(*b*) The dogmatic form of the ' true/false ' test item, with its provocative expression of a point, did seem particularly well adapted, as was expected, to stimulate discussion. Against this it might be objected that that form is appropriate only where the facts, falling naturally into a simple dichotomy, point unmistakably in one direction or the other. In other words, what is gained in heightened interest may be lost several times over through distortion of the important issues. The danger clearly exists, but it was also clear on this occasion that this feature of the ' true/false ' form may be turned to account by pointing out in discussion that such-and-such an item is an unsatisfactory account of that particular question precisely because it states it in the form of a false opposition.

GENERAL APPLICATION.

The procedure described above was used once only, but it might reasonably be expected to have fairly general applicability since, as suggested already, it seems to operate according to the general educational principle that it is important to make self-esteem and social approval dependent on continuous interest and participation. In other words, its efficacy seems to depend upon the ' ego-involvement ' of each group member being formally secured.

Another aspect of this question of generality is whether other forms of ' new-type ' test (e.g., the multiple-choice or re-arrangement types) would be equally or even more effective for this purpose than the ' true/false ' form. Naturally, any sort of test, including the traditional essay-type examination, yields data with this potentiality as one of its by-products, but new-type tests seem particularly well adapted to laying bare the opinion-structure of a group.

Indeed, a test like that used on this occasion may be regarded as a multiple attitude-survey which, though extremely rudimentary, can be sensitive enough to give valuable insight into the pattern of opinions in a group. On this occasion there was no intention to use that insight to exercise any very strict control over the course of discussion, apart from what happened in the case of item 3. That exception shows clearly, however, that the test had a reconnaissance value which helped considerably in the choice of battle-ground and the preparation of positions to defend.

The possibility of manipulating a discussion in this way prompts the question : What exactly is a good discussion ? In practice, it is rare for the objects of a discussion group to be precisely, or in any way explicitly, defined and it can sometimes be sensed that the objects of leader and members are not in perfect harmony. However, whether a discussion's objects are tacit or explicit, they imply value-judgments, and these latent value-judgments require inspection from time to time. The procedure described above tends to force attention on to the general and particular aims of a discussion, thereby subjecting its underlying assumptions to a scrutiny and clarification which can indirectly do much to improve its quality.

There are several obvious obstacles to the extensive use of ' new-type ' tests as reported here. The construction of the test is an arduous undertaking, but if such a test were needed in any case for ordinary measurement purposes, this would not be a difficulty. Also, an interval between giving the test and holding the discussion is necessary for scoring and the analysis of results, and for any special preparation which may then be indicated. However, difficulties of this sort are not insuperable : the crucial point is whether the results achieved are obtainable with less trouble by more established means. The decision reached in any particular set of conditions must depend ultimately on such factors as the sort of topic to be discussed, how much insight the discussion leader has already into the group's attitudes to that topic, and his views on his ability to run a profitable discussion without such pettifogging preparation.

ACKNOWLEDGMENTS.—This exploratory study was made in co-operation with Professor C. A. Mace and Miss P. M. F. Hooper, and their resourcefulness and enthusiasm are gratefully acknowledged. Thanks are also due to the members of the discussion group, and to the Principal and other members of Brixton School of Building who enabled it to be carried out in the course of the School's activities.

CORRESPONDENCE

On S. WISEMAN'S Article : " TRENDS IN EDUCATIONAL PSYCHOLOGY "

Although Stephen Wiseman produces a number of possible reasons for the deficiencies of educational research in this country, he fails to mention a very curious feature of the situation, and that is the very meagre output of the people who are called ' educational psychologists.' According to the B.P.S. Annual Report for 58/59, there are 204 members of the Children's Section of the English Division of Professional Psychologists (Mental Health). Some of these will work in the N.H.S. and could call themselves ' clinical psychologists ' but most of the others will work for local education authorities and are, properly speaking, the only ' educational psychologists ' in the country. It is of some interest to consider the failure of these people to do research, since Wiseman's review seems to be concerned only with (*a*) those temporarily engaged in educational psychology, i.e., teachers working for higher degrees—and (*b*) those employed by institutes and training colleges to teach the subject and (sometimes) to do research. The unique feature of the ' educational psychologist ' is that he is a professional psychologist working full-time with children as a psychologist. Surely *he* is best fitted by training, experience and opportunity to produce educational research ?

Some reasons for his failure are obvious :

(*a*) He is very much over-worked by the local education authority and commonly has no time to spare for research. Wiseman may be alarmed at the small number of teachers going on to Ph.Ds., but I think we should be surprised at the smaller numbers of educational psychologists who are able to work for a higher degree. Offhand, I can remember only one in recent years who has managed the work for a Ph.D., Tyerman, the Educational Psychologist for Grimsby. In my own case, I began work for a Ph.D. one year ago and have made negligible progress. In one's everyday work as an educational psychologist, it is almost impossible to collect data which would be useful for research, despite the fact that the *problems* with which one is faced cry out for investigation. It is this close acquaintance with problems which causes educational psychologists sometimes to reject university-based studies which may fail to satisfy their demands for help and knowledge.

(*b*) Employers—the local education authorities—are by no means always sympathetic to the idea of research. Their attitude is often one of " You are or should be so busy that you have no time for *that*." I have had my own efforts to do research vetoed in this way. Furthermore, research costs money. If the local education authority refuses to foot the bill, the educational psychologist is not likely to. Not unnaturally, we find little sympathy amongst councillors.

But lack of time and local resistance do not entirely account for the educational psychologists' silence. Probably the more profound and general reason is to be found in the educational psychologists themselves. How many of them are fitted by interest and training to produce research ? In the absence of a survey of the experience and qualifications of educational psychologists we can only guess the answer to this as " Not many." Their profession is a very young one and many of them must have become involved in the work almost incidentally, certainly without being hagridden by professional courses and the necessity for qualifications. It would be of value to know how many educational psychologists do not have degrees in psychology. As recently as 1955, the Underwood Committee was able to discuss the possibility of dropping the psychology degree as part qualification of the educational psychologist, and it is still possible for a graduate in any subject whatsoever to become an educational psychologist by taking a one-year course in general psychology (and sitting the M.A. Qualifying Exam.) followed by a year's clinical experience. Unless such recruits to the profession are of outstanding quality they are not likely to merit the description ' psychologist,' nor to be interested or competent in the mounting of research projects.

The failure of training college lecturers and educational psychologists to do research is linked with the lack of interest and plain ignorance of teachers. Again we lack a survey, but my guess is that few training colleges employ a lecturer who is competent to teach and to practise educational psychology. In earlier but post-war years, I went through a two-year training course and received lectures in educational psychology. These were given by the lecturer in physics and chemistry and were sincere but useless. My passions aroused by this experience, I have scrutinized the jobs columns of the *Times Educational Supplement* for years but can hardly recall a training college advertising for a Lecturer in Educational Psychology. I think a survey on the lines of Hearnshaw's "Medical Curriculum" (*B.P.S. Bulletin*, May, 1959) would be both invaluable and salutary.

We therefore have a nice vicious circle. Teacher students are left in ignorance of, or actively turned against, the uses of psychology in education—they tend to be antagonistic towards the educational psychologists whom they meet in the local education authority—the educational psychologists are often not able to deal with the teachers' misconceptions because of their own deficiencies in experience and training—the teachers are confirmed in their misguided notions. This could only be broken by : (*a*) insisting that all training colleges have an educational psychologist on their staffs ; (*b*) by insisting that all educational psychologists in local education authorities have, in addition to their other qualifications, at least a degree in psychology ; and— if it were possible—(*c*) by insisting that part of the work of an educational psychologist is research. (*a*) is very feasible and ought to have the effect, apart from anything else, of turning the training college into a research and advisory centre for its district, not only for its teachers but also for its children ; (*b*) obviously arouses personal antagonisms and whilst essential is likely to have difficulties. I am very pessimistic about (*c*).

E. C. HOPKINS.

. . . . . .

I should like to comment on certain features of Dr. Wiseman's interesting paper on research in education.

It may be of value if we first glance at some research problems in other disciplines.

In medicine, as in many other sciences, the workers in fundamental research are usually members of a team in a particular institution of university rank. The user—the general practitioner in the country surgery—is too busy and too isolated to be able to deal with problems of this nature. Yet, on the other hand, he has extensive records over many years covering many aspects of his patients, and has the material which, if he had the leisure, would enable the family doctor to conduct investigations of value both for his own practical day-to-day work and for other related disciplines such as those of demography and of sociology.

Again, in the field of applied statistics, the reader of the appropriate journals will note that the majority of the research articles seem to come from the individual out in the field, the lone statistician in the works, though much of the fundamental research is, again, of course, being done by members of particular universities.

When we turn to education, the practical man who leaves the opportunity of working on fundamental research goes out into the field to his post as a teacher. Unlike the applied statistician, he will find that he has no junior staff, of computor and typist, to compile his figures and do his preliminary analysis. The teacher has to compile his own figures, just as the doctor has to compile his own individual case histories. He is probably not invited to report to any central body about any research work that he may have in progress.

As an illustrative example, let us consider one feature of the teacher's material—his mark book. Here, like the medico, the teacher should be able to contribute material, not merely to his own discipline but also to sociological ones. If his figures are suitable, then, as he has such long-term knowledge of his pupils, he could contribute material for dealing with problems of intellectual growth, of learning, of problem-solving, of teaching methods in specific cases, of class relationships, as well as for helping with such sociological problems as those relating to gifted families ; to ability and age ; to ability, family size, inbreeding and differential fertility ; to natural ability and social or economic class ; to ability and locality. But all this depends on his figures being suitable, so that much investigation awaits a proper method of keeping his returns of school marks. What are the comparative reliabilities and validities of his marks, in homework, in classwork, in tests (the teacher will continue to use his own tests, set on the specific syllabus he is working to and fitted to the class that he is working with, in preference to any externally prepared tests), and in examinations in the various subjects of the curriculum, and how can they be compared and combined for various purposes ?

This is, perhaps, not fundamental educational research : it is, however, very necessary for the practical man. But what can be done by the lone worker set in a school remote from any university, unable to afford or to house all the possibly relevant literature. If he does any investigation in such problems as instanced, he may find (as I have done) that the educational journals will regard his paper as being too much of a statistical nature and the statistical journals as being too much of an administrative nature. As Dr. Wiseman points out, his own professional bodies will not publish such matter, for too many of the executive committees of such bodies are still to be persuaded that research of any kind in their field is of any value (naturally, perhaps, for they know how worthless much so-called research is when done by the young student who is, in their opinion, ignorant of the very first thing about the craft or about the child). Every educational scientist who has sat in at staff meetings has heard the head and colleagues refer to their marks (often in a very unreliable terminal examination) for decisions, on promotions, repeating or demotion, on prize allocation, on teacher's assessments, on stream, form and set grouping, or even for entry on cumulative record cards or reports. All this is done without the teacher even realising the elementary fallacies of so-called percentages, let alone the necessity for comparability by similar mark frequency distributions. The critic will appreciate that a proper method of handling school marks is essential for the good conduct of any school, let alone for the investigation of the many questions—such as those suggested above—where the teacher in the field has the material to throw light on so many educational and social problems.

I suggest, therefore, that Dr. Wiseman should not be surprised that so few teachers conduct any research once they leave their post-graduate course. On the other hand, the majority of the problems of education that worry the practical man are not those that can be solved by an investigator who is not a member of the school community but who comes in for a few *ad hoc* sessions to get some nebulous statistical pictures. Rather they can be dealt with conveniently by the man on the spot who has the fruits of the work with his children over long periods of time, who knows in each case the whole child and many of his special circumstances, and who knows the fallacies that can beset any superficial interpretation of his data.

FRANK SANDON.
(*Sometime Headmaster, Millom School, Cumberland*).

REPLY

Mr. Hopkins raises a matter the importance of which I do not dispute. My only excuse for not referring to it in my paper is that I did not feel it desirable to include it unless it were accompanied by a fairly full discussion of the special circumstances surrounding educational psychologists working in clinics and the National Health Service—and the article was too long as it was. There is, I think, little doubt about the ' meagre output ' of research of those who, in Mr. Hopkins' phrase, " are, properly speaking, the only ' educational psychologists ' in the country." (In passing, I could agree whole-heartedly with this quotation if ' prejudicially ' were substituted for ' properly ' !). I doubt whether many research workers in educational psychology would agree that clinic psychologists are " best fitted by training, experience and opportunity to produce educational research." Indeed, Mr. Hopkins himself is not too sure in view of the third sentence of his third paragraph. But I would certainly agree that they are best fitted to pursue *some* of the problems of educational psychology, and it is unfortunate that by inclination (for some) and by pressure of work (for many) their output is so small. Mr. Hopkins suggests three ways in which his ' vicious circle ' might be broken. His (*a*) is an impossibility for many years to come, because of the supply position. Some colleges *have* educational psychologists on their staffs : many more would like to have them. His (*b*) is not so simple as it sounds : I would be reluctant to see an absolute rule here, since *some* of those who come up the ' non-conformist ' way are more orientated towards research than their confrerés. With (*c*) I am in complete agreement.

Mr. Sandon is not, I fear, fully aware of the present position of teachers and research. Nearly all the research done at the Master's level is done part-time by practising teachers. They *are* members of the school community, often working on material provided by their own classes. Indeed, my main argument about Ph.D. regulations is that they demand full-time research, whereas a longer period spent part-time may prove much more valuable in attacking real problems. I agree entirely with Mr. Sandon about the difficulties of " the lone worker set in a school remote from any university." But this is true of the isolate in any research field. Mr. Sandon is, I hope, putting forward Carlisle's claim as one of the new university towns.

S. WISEMAN.

An Investigation into the Response of Boys and Girls Respectively Towards the Content of the History Syllabus in Grammar Schools.

BY IEUAN JONES

(*Summary of a thesis presented for the degree of M.A. at the University of Wales*, 1955).

DESIGN.

A Questionnaire was given to the fourth forms of four boys', four girls,' and four mixed grammar schools in South Wales. There were 190 girls and 143 boys in the tested population. The Questionnaire concerned pupils' response to :

(1) History as a school subject.
(2) Six main topic groups—Military, Religious, Social, Cultural, Economic, Political (each of the groups being divided into four sub-sections and illustrated by examples).
(3) Contemporary History.

The interest in these topics was recorded on a five-point scale ; there was also provision for free response. The replies were anonymous.

THE RESULTS.

(1) *History as a school subject.*

A high proportion of the pupils (87 per cent. of girls and 85 per cent. of boys) expressed a very great liking, or liking. Only three (all boys) expressed ' a very great dislike'. In their free responses the boys showed a tendency to refer more often than girls to military, political and economic aspects, and the girls more often to social aspects.

(2) *Six Main Topic Groups.*

A more detailed examination was then made of the response towards the various aspects of the syllabus in history. There were six main topic groups, each of which was divided into four sub-sections—there were, therefore, twenty-four items in all. The response towards each item was recorded on a five-point scale (+10 +5 0 −5 −10). When the scores on all twenty-four items were combined, the following averages were obtained : All Pupils (79·70) ; girls (84·50) ; boys (73·32). Thus here, where the maximum possible range was +240 to −240, this difference of only 11·18 in favour of the girls may be described as being rather slight. This difference in means was, in fact, not found to be statistically significant.

Turning to an examination of the response towards each of the six main topic groups, where the maximum possible range is +40 to −40, we may observe the manner in which this over-all 11·18 average difference in favour of the girls is made up :

TABLE 1

RANK ORDER OF DIFFERENCE IN MEANS.

	Topic	Girls' Av.	Boys' Av.	Difference.
1st	Economic	5·61 (6th)	17·59 (2nd)	11·98 (B)
2nd	Religious......	14·53 (4th)	3·74 (5th)	10·79 (G)
3rd	Social	21·89 (1st)	13·99 (4th)	7·90 (G)
4th	Cultural	9·50 (5th)	2·55 (6th)	6·95 (G)
5th	Military	17·03 (2nd)	19·72 (1st)	2·69 (B)
6th	Political	15·95 (3rd)	15·73 (3rd)	0·22 (G)

This table reveals that marked rank order differences occur only in two groups—Economic and Social, and that the greatest score differences between the sexes arose in the Economic and Religious Groups.

Using the Lindquist 'small sample error' formula to test the statistical significance of the difference in means between the sexes for the Six Topic Groups, four were found to be significant (Economic, Religious, Social and Cultural), while the other two (Military and Political) were not. The amount of the difference, even where significant, was however small in relation to the total range of the scale; this was particularly so with Cultural and Social History.

Apart from these comparisons based on attitude scores towards twenty-four items, the pupils were later asked to place the six topic groups in an order of preference—this would provide something of a reliability check on the first conclusions. (The Method of scoring used allocated 6 points to a topic if it was placed first in order of preference, five points if it was placed second, and so on down to 1 point if it was placed sixth—last). The result was as follows:

TABLE 2

Topic	Boys' and Girls' Score	Girls' Score	Boys' Score
Military	2nd (1395)	3rd (691)	1st (704)
Religious	3rd (1144)	2nd (752)	5th (392)
Social	1st (1536)	1st (953)	2nd (583)
Cultural	6th (910)	5th (560)	6th (350)
Economic	5th (952)	6th (439)	3rd (513)
Political	4th (910)	4th (595)	4th (461)

The chief features of this Table appear to be:

(i) Social and Military history emerge as clearly the first choice of girls and boys, respectively—this agrees with the results ih both Table I and Table 3.

(ii) The Religious aspect (with the Economic aspect), provides the biggest rank order difference between the sexes—in Table 1 it accounted for the second largest score difference between them.

(iii) Social History occurs first and Cultural aspects last in the rank order for all pupils combined—in both Table 1 and Table 3 they were second and last, respectively.

(iv) The over-all response pattern of the sexes emerges as generally similar apart from the Economic and Religious History Groups—a parallel comment was made in Table 1. It is also interesting to note that the two topics least liked by girls and boys, respectively, remain the same in all three Tables. (In Table 3, the rise with both sexes in the position of the Political aspect may have resulted from the biographical form in which the topic was in this case set down.)

(3) *Contemporary History.*

80 per cent. of the girls and 86 per cent. of the boys recorded that they would like to have lessons on contemporary history, over 40 per cent. of the girls and over 50 per cent. of the boys expressing a response of ' would like very much.' As in the case of History as a school subject, the boys in their free responses tended to refer more frequently to such items as war, science, industry, and politics, and the girls to refer more often to social aspects.

In order to fill out the information concerning the response towards Contemporary History, the pupils were required to amplify their reaction by arranging six Contemporary History Topics in what would be their order of preference. The rank orders were as follows :

TABLE 3

Topic	All Pupils		Girls		Boys	
Military	1st	(1479)	3rd	(759)	1st	(720)
Religious	5th	(890)	4th	(606)	6th	(284)
Social	2nd	(1476)	1st	(894)	3rd	(582)
Cultural	6th	(825)	5th	(532)	5th	(293)
Economic	4th	(946)	6th	(439)	4th	(507)
Political	3rd	(1377)	2nd	(760)	2nd	(617)

COMMENTS.

(1) Pupils likes and dislikes should not be allowed to determine the content of the history syllabus, but ought to be allowed, in the interests of more efficient teaching and learning, to colour the teacher's approach.

(2) References to contemporary history almost invariably contained references to the ' worth-whileness,' usefulness and interest of the material. There is obviously a powerful source of interest here which it would be both wasteful and harmful to ignore.

(3) No marked differentiation between the sexes in the teaching of history seems to be required as far as the tested population was concerned, although it was noted that certain aspects were differently favoured by boys and girls, while the free responses revealed certain attitude variations and differences of tone between the sexes which suggested that the two sexes often tend to approach an aspect from a different angle.

(4) It is necessary to stress the response similarities between the sexes which were revealed ; important individual divergencies of interest revealed a variability within each sex which was more important than that between the sexes. The prime consideration seems to be to cater for individual pupil variation rather than for differentiation between the sexes.

(5) While it is not always possible to please equally both sexes, there would appear to be little in the syllabus which is absolutely distasteful to either. The differences which were observed—particularly as it was noted they were not of a pronounced nature—could be used as an argument both for and against the co-educational teaching of history. The case in favour would be that the greater enthusiasm of one sex for any particular aspect might possibly spread to the other. This possible transfer of enthusiasm could also result in the historical outlook of both being broadened and the tendency of each sex to direct their attention too exclusively to any particular side of history might be corrected.

(6) As noted by so many researchers on similar themes, the same general reasons for ' liking ' and ' dislike ' appear. The two most frequent are interest and usefulness ; proficiency is hardly ever mentioned, though it may well have a hidden effect.

BOOK REVIEWS

ADLER, A. (1958). *The Education of the Individual.* New York: Philosophical Library, pp. xiii+142, $3.50.

The title of this book suggests that it could be of considerable value to teachers in the understanding either of themselves or of children. Unfortunately, however, the educational thinking that it contains is much lower than that of a good training college discussion group. Apparently the author takes incidents from others as stimuli to free association, and such a technique inevitably acts as a buffer against reaching any valuable material. The main conclusion seems to be that all men are equal because anyone, from moron to genius, may stimulate another to valuable thought. The book reads like a social philosophy which is the compensatory fantasy of a despairing solitary.

There are frequent references to myth cycles, to stories by writers from Virgil to Proust, to characters from Citizen Kane to Helen Keller; and it is clear that the erudite author is following some train of thought which is lucid to himself. But if it is not also clear to the reader, then the book is not ready for publication. The resulting confusion is particularly disappointing to educationists, for a number of reasons. First, there is a crying need for greater insight into motivation and purpose in the handling of young children; but in this book analogy and poetic prose supplant the necessary logic. Then there is need for increased precision of observation and statement; but in this book mysticism clouds sincerity. P. M. PICKARD.

COLLIER, K. G. (1959). *The Social Purpose of Education.* London: Routledge and Kegan Paul, pp. xv+236, 21s.

This book is a successful synthesis of educational psychology, educational sociology and philosophy, at a moderate level of difficulty. It is intended mainly for the teacher and student teacher and on the whole it hits the target, though some parts would be rather difficult for the less able training college student. It should also be read by many educationists, and though psychologists will learn no psychology from it, they might broaden their horizons and benefit from studying the neat way in which these several aspects of education are linked together. So many books on the philosophy of education merely tell us, with modifications, what Plato told us several thousand years ago; they have some use in reminding us of ultimate values, but they advance education by not one iota. Though the author naturally ends with the same values, we are interested in the path by which he arrives there.

The book is concerned with the problems of education in a rapidly changing society, and is divided into four sections: The Challenge of the Present Situation; The Response—Aims and Resources; The Underlying Assumptions; Educational Principles. Planted with his feet firmly in the school (albeit one which is a trifle idealistic), the writer discusses the techniques which should be used in order that pupil and teacher should achieve the philosophic aims of society. In particular, he argues the principles of discipline, pupil participation and the instructional methods needed for this purpose. One of the best parts of the book is the preface, which should be read by all critics of university departments of education. The style is clear and concise and the book is good value for the price. The third table, on page six, is badly presented (the right-hand figures should be confined to group totals).

R. R. DALE.

GALLAGHER, J. R., and HARRIS, H. I. (1958). *Emotional Problems of Adolescents.* New York: Oxford University Press, pp. 174, 28s.

This is a book written by two optimistic and warm-hearted American doctors and it is addressed to all who deal with adolescents in groups or individually. It emphasises that each adolescent is a person, not a problem, and that the role of the adult towards him is not so much to protect him against adversity, as to understand him, guide him and give him temporary assistance. There are useful reminders that an adolescent may try to identify himself with one of his parents and about the need to let adolescents talk out their feelings instead of acting them out in anti-social behaviour.

The book is definitely for non-specialists; many of its statements will seem to be common knowledge to the expert. But it is the sort of book which, because of its understanding and friendly attitude to the problems of adolescents and of the adults who are helping them, can do little harm and could do quite a lot of good. It will certainly encourage parents and teachers to stop and analyse their own feelings before guiding the adolescents in their charge. M. METCALFE.

GARRETT, H. E. (1958). *Statistics in Psychology and Education* (Fifth Edition). New York and London: Longmans, Green, pp. xii+478, 35s.

SENDERS, V. L. (1958). *Measurement and Statistics.* New York: Oxford University Press, pp. xvi+594, 48s.

DOWNIE, N. M. (1958). *Fundamentals of Measurement: Techniques and Practices.* New York: Oxford University Press, pp. xi+413, 48s.

The reviewer has found Garrett's textbook so clear and useful, for over thirty years, that he sometimes wonders why so many books in the same field are published. This fifth edition is re-set to be easier on the eye, but does not include much new material beyond the fourth. A few additions are the effect of heterogeneity on correlations, some of the simpler non-parametric tests, the use of stanines, the cosine-pi technique for tetrachorics, and a fuller treatment of test theory.

Miss Senders, however, has accomplished the seemingly impossible in writing an original book on her topic. It is structured firmly around the distinction between nominal, ordinal and interval scales of measurement; and the introduction of some elementary concepts of information theory at an early stage makes it possible to approach correlation as a type of reduction of uncertainty, i.e., through analysis of variance. Also, later, there is a particularly thorough treatment of non-parametric techniques. Less happily, the book sounds like a direct transcript of the author's lecture courses. Thus, if one wanted to use it for reference purposes, one would have to plough through such passages as: " If you think that people who are good in one subject tend to be good in the other, you may conclude that since Arbuthnot was better than Zilch in psych, he was probably better in stat too, and you'd guess that he got an A or B in that subject." It would be an excellent book for the student who wished to teach himself statistics without attending any lectures, provided he could put up with this chattiness and with being told everything five times over.

Downie's book is misnamed: it has little to do with fundamentals. Rather, it is a sound and useful textbook for an elementary course on intelligence, attainment and personality testing. It is particularly directed to teachers and school or college counsellors, and is extremely practical in its advice throughout. Three chapters are given to the simpler psychometric techniques, and there is a lengthy section on constructing new-type tests. For British readers perhaps the most useful portion is the summary descriptions of some of the better-known American tests. P.E.V.

GODDARD, N. L. (1958). *Reading in the Modern Infants' School.* University of London Press, pp. 142, 13s. 6d.

This book is not an account of research work carried out, nor is research mentioned other than in the most general way. It is rather a handbook of method for practising teachers, stating the general principles and detailing ample suggestions for their implementation in the classroom.

One basic principle stated is that all children should be introduced to reading through whole sentences arising from their interests and written down in individual or class news-books, or in a wall story. The need for supplementary exercises is recognized, but it is suggested that this approach generally should be maintained until children begin to progress in reading. Another stated principle is that the active promotion of phonic and word analysis should not take place until the middle of the second year in the infants' schools.

Whether accepting these principles or not, all teachers would benefit from considering the many valuable suggestions for using interests arising naturally from other classroom activities as sources of inspiration for reading material and reading activities. And the plan outlined for organizing the more formal teaching of reading seems a highly efficient method for the distribution of the teacher's time and effort in a class with groups of varying reading ability.

In general, the book is well-considered and balanced in its outlook and emphasis. It is felt, however, that Miss Goddard too readily accepts the fact that some children will make little or no progress in reading during their stay in the infants' school. Difficulties exist, but is it generally an unattainable goal for teachers, given reasonable co-operation, to have all—not some—children reading fluently and well before transfer to the junior school? Might not this happy state be more readily achieved if flexibility of approach, which Miss Goddard recognizes as generally desirable, were extended to the beginning stages, so that children failing to make progress in reading, with the writing and reading of whole sentences about topics of interest, are given, not more experiences of the same kind, but other approaches, varied according to the interests and abilities of both teacher and child?

E. J. PEARSON.

HALMOS, P., and ILIFFE, A. (edit.) (1959). *Readings in General Psychology*. Routledge and Kegan Paul, pp. ix+251, 25s.

The Editors, in bringing together some fifteen papers from fourteen contributors, including themselves, aim to provide for the student of psychology a textbook which would cover a limited number of topics. These selected topics, however, were intended to be dealt with in greater detail than is customarily done in more general introductory textbooks. The Editors do not claim that this book constitutes a complete introductory course; nor that the subjects chosen either adequately cover the field of psychology or are necessarily the most important topics or issues in that field. This approach, it is claimed, is more satisfactory for the student with limited time for training and avoids the usual hasty skirmish with too many topics which does not allow the student to reach a thorough grasp of any.

The papers contributed tend to be somewhat on the lines of the popular 'readers digest' type of article—quite readable, but lacking any continuity between one paper and the next and in some cases (for example Osselton's paper on electroencephalography) appear to assume too great an existing knowledge of the subject by the reader.

The selected references for further reading, given at the end of each paper, should enable a student to build up his knowledge of the topic presented in the paper. Some dozen pages at the end of the book are devoted to references, which, surprisingly, do not include the names of the publishers of books listed.

It is difficult to see this book used as a textbook, but it will be useful as supplementary reading for students who have gone some way through an elementary course in psychology. It will also be helpful to more advanced students or to more general readers (such as school teachers and others interested in social work) who wish to acquire quickly a reasonable working knowledge of some of the particular topics selected by the Editors.

W. ROBERTSON.

MCFARLAND, H. S. N. (1958). *Psychology and Teaching*. London: Harrap, pp. 232, 10s. 6d.

The titles of the fifteen short chapters of this textbook of educational psychology are very similar to those in most recent ones on the subject. They deal with such topics as learning, child development, psychological assessment, personality and mental health. Each chapter is a synopsis of modern psychological thought on its particular field presented in an orthodox, moderate manner, giving weight to divergent views and not disguising the inability of psychology to provide answers to all educational problems. The language is usually simple and lucid and the immediate applications of general principles to classroom practice are pointed out. Not all the chapters are equally successful. However there are, for example, good ones on Discipline, Delinquency, and Deprivation, and on Aesthetic and Technical Education. These probably give the practising teacher most of the general information he needs on these topics. Others are so superficial or condensed that they make little impact. This is specially true of the chapters on intelligence and on educational assessment. The fact is that these topics are far too wide to be compressed into so few words while yet retaining a readable style ; further, some of the concepts are so novel to many elementary readers that they need to be presented in a number of different contexts before they can be really used as media of communication. The chapter on assessment, for example, deals with such topics as variance, standard scores, correlation, reliability, validity and factorial analysis, all in sixteen small pages.

All this raises the question : " For whom is the book written ? " The author indicates that it is for students in training colleges. For this purpose it will no doubt, find a place, but not as *the* textbook for the course. Each chapter may prove a useful starting point to a series of lectures, discussions and guided reading of more specialised texts ; the book can in no way be regarded as the main basis of study. For the more mature students in university departments of education and for teachers who, some time after they have left college, are about to embark on an extended course of study, perhaps leading to a higher qualification, this little book may prove valuable as a means of filling in the general background and defining the field of study before beginning the intensive work. L. B. BIRCH.

McINTOSH, D. M. (1959). *Educational Guidance and the Pool of Ability.* University of London Press, pp. 199, 21s.

With the help of a large panel of volunteer teachers, Dr. McIntosh has largely repeated and extended the previous study of selection for secondary education which McClelland published in 1942. The differences are interesting : he has followed up a procedure which is in everyday use instead of an experimental one ; he has compared its results with pupils' progress and examination success throughout their secondary careers instead of with third year ratings only ; and he has covered a whole county age group (Fife) instead of a single city (Dundee). Notwithstanding, the validity correlations are very similar to those previously reported and, when corrected for restriction of range, they remain as high over 6 years as over 3. The combination of two I.Qs., semi-objective tests of attainment and scaled teachers' estimates in English and arithmetic correlate ·872 with all-round marks in the fifth and sixth years of the senior secondary school. Perhaps the most novel finding is the superior validity of teachers' estimates and the test of ability in English, even for Maths. and Science marks, and the lower validities of the arithmetic test and estimates ; whereas in the junior secondary schools, with their practically biased courses, the position is reversed. In both types of school the I.Qs. seem to be of intermediate value.

Dr. McIntosh discusses sympathetically the problems of educational guidance, of meeting parents' wishes, of the over-and under-achiever, and of increasing the production of pupils of university calibre ; but he seems to have no new conclusions to offer. The actual proportion achieving good Leaving Certificates in his age group of 4,000 odd is 6·3 per cent., and he estimates that with ideal allocation and parental support, the figure might rise to 11%. But in fact the data indicate the position only when, as in Fife, some 24 per cent. are offered full secondary education ; and they cannot show how many more might achieve the same standards with a different type of secondary school organization. For the same reason, his careful analysis of the ability level required for successful achievement (he concludes that there are even chances of a good Certificate at a mean quotient of 117) might be misleading under rather different conditions. However, this objection to his interpretation of his results in no way detracts from one's appreciation of the value of this extremely thorough and large-scale investigation. P.E.V.

MILLER, D. R., and SWANSON, G. E. (1958). *The Changing American Parent.* London: Chapman and Hall, pp. xiv+302, 52s.

American society, as several writers have recently pointed out, is in transition from go-getting individualism, with its sustaining myth of every schoolboy a potential millionaire, to corporate organization which stresses, not raw individualism, but the values of loyalty and common purpose. Competition is now tribal rather than individual; no longer Smith *versus* Brown, but Smith, Inc., in competition with Brown, Inc. How does this affect the aspirations of American parents for their children? In what ways is child-rearing being modified in order to promote the attitudes and values of personal life now in demand? What are likely to be the ultimate effects of these changes in parental expectations and methods? These are the questions explored in this book.

The authors based their work on a sample of 1,157 families in the Detroit area. They abstracted a traditionally orientated group (individuated-entrepreneurial !) defined by the closeness of the parents' association with individualist private enterprise; and an emergent group (welfare-bureaucratic) defined by socio-economic dependence on large-scale enterprises. It was assumed that the traditional families would train their children in ways likely to inculcate self-control, self-denial and self-reliance through the internalization of a strict code of values, while the emergent families would be more likely to engender adaptive values, appropriate to companionable co-operation and here-and-now enjoyment of living. These differences, it was predicted, would best be shown in the parents' attitudes to weaning, bowel and urinary training, the child's pleasure in its own body, rewards, punishments, etc.

A number of fruitful indicators emerged at levels of significance which suggest that the authors are on the track of a social index of great value. Those concerned with social change, the origin of social values, and the relationship of social values to methods of child-rearing will find much to interest them in this book, even though its conclusions are, inevitably, tentative and speculative. In addition to its research content, the book includes a fascinating resumé of the dominant ideas in training American children from 1,800 to the present time and a good deal of useful general information on social trends in the United States.

J. HEMMING.

MILLER, K. M., and NICOL, M. A. (1958). *Occupational Status, Sex and Age as Factors in Radio Programme Choice.* University of Tasmania, Department of Psychology, Publication No. 1, pp. 79, no price given.

One of the newest Psychology Departments in the Commonwealth, that of the University of Tasmania, has now embarked on its own series of research publications. The present issue surveys the radio preferences of some 600 adults interviewed at their homes in Hobart. The work was done by eighteen students and staff members, and must have been a useful exercise in applied psychology. The results, broken down by occupational grade, sex and age show much the same patterns as are obtained in this country.

MORRIS, JOYCE M. (1959). *Reading in the Primary School*, with a Statistical Appendix by P. M. Grundy. National Foundation for Educational Research, Publication No. 12. London: Newnes, pp. xviii+179, 15s.

In 1953 the National Foundation for Educational Research was invited by the Kent Local Education Authority to enquire into the abilities and attainments of children in Kent primary schools. This book contains a description of the first part of the enquiry, relating to reading performance in a random sample of sixty schools, containing over 8,000 children, in urban and rural areas of varying socio-economic status. Children of all ages were given the Watts Sentence Reading test; and the 11-year-olds were given the Watts-Vernon Reading test. Three tests of non-verbal intellectual ability were used, appropriate for successive age groups. In addition, a careful study was made of the different types of school, their size, organization, type of building, size of class, and the methods by which the children had been taught to read.

It was found in the first place that the average reading achievement of the 11-year-olds in the Kent schools was significantly superior to the national average as obtained in the Ministry of Education survey in 1954. There was a significantly higher percentage in the class of superior readers, and about the same proportion of illiterates and semi-literates, the actual number being very small. However, an estimate based on the reading books used showed that nearly half the children in the first year at the junior school still required the kind of teaching associated with the infant school. But the author has postponed until a later publication any discussion of the type and degree of reading backwardness in these schools.

It was rather unfortunate that it was thought necessary to make this postponement, and to fill more than half the present volume with descriptions of school organization and methods, and teachers' opinions of these. The latter are in the main familiar enough, and not particularly relevant to actual experimental findings in relation to reading.

However, a careful statistical study, the methods and detailed results of which are described by Dr. Grundy in the Appendix, was carried out of the association between reading test performance and a number of characteristics of the schools and their teaching. As may be expected, the associations were complex, since there was much interaction between the various factors. It was found that although average reading performance was better in larger schools, with better buildings, in urban districts of higher socio-economic status, the average intelligence test results were also higher, and this seems to have accounted for the differences in reading. Allowing for intellectual capacity, superiority in reading performance appeared in junior schools not associated with infant schools ; and in children who had received a considerable amount of phonic teaching in their infant schools. The latter result appeared only when adjustment was made for the intellectual inferiority of the children who had been taught in this way. Neither of these findings is readily explicable ; and it may be that they are fortuitous, or due to other causes not yet apparent.

As a general investigation of the factors making for good or poor reading, the study described here is disappointing. We must hope that the further more detailed studies of these to be published later will throw more light on the problem.

M. D. VERNON.

PARKYN, G. W. (1959). *Success and Failure at the University*, Vol. I. Wellington : New Zealand Council for Educational Research, pp. xvi+254, 35s.

In New Zealand, about one-eighth of each age group reaches one of the four universities, and of these one-third of the full-time and over half the part-time students fail their courses at first-year standard (Stage I). The Senate of the combined universities invited Mr. Parkyn, the Director of the N.Z. Council for Educational Research to investigate the situation and to recommend any steps that might be taken to reduce student failure. His enquiry encountered the usual difficulties of handling school and university gradings, but he received exceptionally full facilities and is thus able to present an exceptionally thorough analysis of the qualifications and performance of over 4,000 students. The overall correlation between school performance and university success reaches much the same level as has been found elsewhere, corresponding to a coefficient of about 0·54, though it is much lower for part-time students and for second or later year students taking Stage I subjects. Perhaps the main points of interest for English readers are :

(*a*) The great majority of students enter as a result of ' accrediting ' by their schools, not, on the basis of an external examination. The schools in general maintain high standards and such students are by no means inferior to those admitted by formal examination.

(*b*) The author provides a particularly clear discussion, for the benefit of statistically untrained readers, of the nature and implications of correlation, and shows how raising of entry standards would, under New Zealand conditions. cut out far too many successful students.

(*c*) He has devised some ingenious techniques for dealing with such awkward predictors as sixth-form rank orders, and such criteria as course grades. Interesting also is his use of Kendall's tau coefficient with contingency tables as a non-parametric measure of correlation.

A second volume will present, next year, the results of Mr. Parkyn's studies of causes of student failure.

TRAVERS, R. M. W. (1958). *An Introduction to Educational Research.* New York: Macmillan, pp. xviii+466.

Though Dr. Travers is a product of this country, he has been in America long enough (he is now Chairman of the Educational Psychology Department at the University of Utah) to have absorbed completely the verbosity and sophistication which characterise so much American educational writing. A general textbook for Higher Degree students on the principles and practice of educational research would be invaluable; but this one is too hard going for any except Ph.Ds. One has to struggle through some 150 pages, mainly concerned with the rationale of the scientific approach to education, before one reaches the more useful sections such as those dealing with validation; observational, rating and interview techniques; survey procedures; predictive studies; longitudinal studies; experimental studies and design; 'processing' research data and report-writing. In all these sections, the student or his lecturer will find much good sense, illustrated freely by weaknesses that have occurred in amateur, student and professional researches. Even here the coverage is uneven—for example there is little on reliability, or questionnaires or test design; and more space is given to shrinkage of multiple correlations and to Q-technique than to factor analysis. Presumably, the author's answer would be that he is more concerned to fill in the gaps which are not covered by standard books on psychometrics. The student may well feel discouraged by the author's trenchant criticisms of most of the research techniques he was probably intending to use, but should also be braced by the scientfic integrity and rigour of his approach. Finally, a more accurate title than Educational Research would be Research in Educational Psychology.

WINNICOTT, D. W. (1958). *Collected Papers: Through Paediatrics to Psycho-Analysis.* London: Tavistock Publications, Ltd., pp. 350, 35s.

Although this book is a collection of papers read to scientific audiences such as the British Psycho-Analytic Society, yet with the exception of a few difficult technical chapters in Part III, it may be read with benefit not only by psychoanalysts, but also by others who are concerned with an understanding treatment of children, e.g., educators, medical practitioners, children's officers, and psychologists.

Because Dr. Winnicott has never " cut loose from pediatric practice," he brings to the exposition of psychoanalytic attitudes a rich experience of normal children outside psychoanalysis. He has a delight in observing child behaviour (" For about twenty years . . . in my clinic at the Paddington Green Children's Hospital . . . I have recorded in minute detail the way infants behave in a given situation " ch. 4), an appreciation of everyday social contexts, and a practical wisdom which remind one of others who have served two disciplines (Gesell, Spock, Susan Isaacs) and been able to convey some of the sense of these disciplines to readers without their special training.

Many of the papers are concerned with " the essential conflicts, chiefly unconscious, that belong to health, that simply mean that the child is alive and lively " (ch. 26). It is important to do nothing at all, " except be a friend," while a child goes through unhappy times of jealousy, disappointment, loss of what is loved, not only because the value of the child's personality would be to some extent diminished by missing these experiences, but also because " surely, a most important aim of education should be to enable the child to manage life unaided." In a paper about the anti-social tendency, with its two trends, represented by stealing and destructiveness, Dr. Winnicott advocates not direct psychotherapy, but the provision by home and school of a reliable human attitude which can stand the strain of impulsive behaviour, in itself a sign that the child hopes to provoke something good from the environment (ch. 25). The contribution to psychoanalysis itself in Part III includes an interpretation of Melanie Klein's concepts. The meaning of " the depressive position " is made clear as an achievement by the infant of a stage of emotional development when he can feel concern.

Allowing for known changes in the history of psychoanalysis, one is still surprised that there should be in these papers such a strong faith in the perceptiveness (without benefit of psychoanalysis) of the ordinary person ; one had thought (mistakenly it seems) that this attitude was more prominent in a different school of psychotherapy, and even there only on the assumption that the therapist is present to act as catalyst, so to speak. However, in recounting the success of parents who have cured severe disturbances almost on their own, it is possible that Dr. Winnicott is underestimating the influence of even fleeting contacts with his own serene and shrewd perceptiveness. Yet he would no doubt be the first to attribute to psychoanalysis some part of the deepening of this sensitivity and tolerance which can be traced from 1931 to 1958 in his papers, and which must have led many who heard him to adopt a more understanding attitude to countless children. L. R. BOWYER.

WORLD FEDERATION FOR MENTAL HEALTH (1958). *Growing Up in a Changing World.* London : H. K. Lewis, 15s.

The Tenth meeting of the World Federation for Mental Health was held in Copenhagen in August, 1957, and the papers which are printed here deal in general with the effect of cultural and social changes on the mental health of young people. The concept of a ' changing world ' is ably discussed by Dr. Brock Chisholm in the final paper. The writers represent psychiatric opinion from five continents and the opening by Margaret Mead ensures that the contribution of anthropology is not neglected. Four papers give factual accounts of the working of the United Nations, UNESCO and the W.H.O. in the field of mental health. In the first of these Dr. Lopez-Rey emphasises the need to distinguish the terms maladjustment, mental health, and delinquency. A number of papers deal especially with the problem of adolescence and among these the most unusual is an account by Dr. Tsung-Yi Lin of two types of delinquent youth in Chinese society, which may help to throw some light on the problem formulated by Dr. Lopez-Rey. A symposium on the prevention of delinquency is interesting as showing how world-wide is the interest in this problem and how diverse the suggestions for dealing with it. Two papers on young refugees provide some new ideas. Of the remaining papers those most likely to interest educational psychologists are one describing the work of a schools psychologist in Denmark and one by Dr. Paul Sivadon (U.S.A.) which traces the family history of two schizophrenic men as far as the great grand-parents. The outstanding value of this collection of papers is the exchange of ideas and frequent similarity of view-points from widely differing cultural backgrounds. M. D. WILSON.

PUBLICATIONS RECEIVED

The mention of a book in this list neither implies, nor precludes, later review.

E. R. HILGARD (1958). *Theories of Learning* (2nd ed.). New York : Appleton-Century-Crofts ; London : Methuen, pp. ix + 563, 45s.

B. F. SKINNER (1957). *Verbal Behaviour.* New York : Appleton-Century-Crofts ; London : Methuen, pp. x + 478, 42s.

W. H. Swift (1958). *Trends in Canadian Education.* Toronto : W. J. Gage, pp. 94, no price given.

R. Gille (1959). *Le Test-Film.* Paris : Presses Universitaires de France, pp.viii + 151, Fr. 1200.

F. Kocher (1959). *La Reeducation des Dyslexiques.* Paris : Presses Universitaires de France, pp. vii + 152, Fr. 480.

N. G. McClusky (1958). *Public Schools and Moral Education.* New York : Columbia University Press ; London : Oxford University Press, pp. xiii + 315, 48s.

N. Malcolm (1959). *Dreaming.* Routledge and Kegan Paul, pp. vii + 128, 12s. 6d.

D. M. McIntosh (1959). *Educational Guidance and the Pool of Ability.* University of London Press, pp. 199, 21s.

R. I. Watson (1959). *Psychology of the Child.* New York : John Wiley ; London : Chapman and Hall, pp. ix + 662, 56s.

D. R. Miller and G. E. Swanson (1958). *The Changing American Parent.* New York : John Wiley ; London : Chapman and Hall, pp. xiv + 302, 52s.

F. J. Mott (1959). *The Nature of the Self.* London : Allan Wingate, pp. xvii + 316, 42s.

K. G. Collier (1959). *The Social Purposes of Education.* Routledge and Kegan Paul, pp. xv + 236, 21s.

J. Deese (1958). *The Psychology of Learning.* New York and London : McGraw-Hill, pp. ix + 367.

A. Weaver (1959). *They Steal for Love.* London : Max Parrish, pp. 132, 12s. 6d.

W. C. Olson (1959). *Child Development* (2nd edit.). Boston, Mass. : Heath, xi + 497, price not stated.

R. Heywood (1959). *The Sixth Sense : An Inquiry into Extra-Sensory Perception.* Chatto and Windus, pp. 224, 21s.,

P. A. Leaning (1958). *The Challenge of Cerebral Palsy.* Auckland, N.Z. : Whitcombe and Tombs, pp. 98, no price given.

J. Hambling, *et al* (1959). *The Nature of Stress Disorder.* Conference of the Society for Psychomatic Research. London : Hutchinson, pp. xiv + 298, 25s.

M. I. Rasey (edit.) (1959). *The Nature of Being Human.* Detroit : Wayne State University Press, pp. ix-115, $3.95.

W. M. Cruickshank and G. O. Johnson (edit.) (1958). *Education of Exceptional Children and Youth.* New York : Prentice-Hall ; London : Staples, pp. xiv + 723, 60s.

D. H. Stott (1958). *The Social Adjustment of Children : Manual to the Bristol Social Adjustment Guides.* University of London Press, pp. 46, 25s. Specimen sets of Diagnostic Forms and templates, 9s. 6d.

C. Burt (1959). *A Psychological Study of Typography.* Cambridge University Press pp. xix + 68, 15s.

S. Hook (edit.) (1959). *Psychoanalysis, Scientific Method and Philosophy.* New York University Press, pp. xiii + 370, $5.00.

E. E. Maccoby, T. M. Newcomb, and E. L. Hartley (Eds.) (1959). *Readings in Social Psychology* (3rd edit.). London : Methuen ; New York : Holt (1958)., pp. xi + 674, 55s.

G. W. Parkin (1959). *Success and Failure at the University, Vol. I. Academic Performance and the Entrance Standard.* Wellington, N.Z.: New Zealand Council for Educational Research, pp. xvi + 254, 42s.

C. Launay (1959). *L'Hygiene Mentale de L'Ecolier.* Paris : Presses Universitaires de France, pp. 211, F.700.

B. B. Wassell (1959). *Group Psychoanalysis.* New York : Philosophical Library, pp. xiii + 306, $3.75.

A. B. Feldman (1959). *The Unconscious in History.* New York : Philosophical Library, pp. 269, $4.75.

E. Bergler (1959). *Principles of Self-Damage.* New York : Philosophical Library, pp. xiii + 469, $6.00.

A. Montealegre (1959). *Formation de la Méthode Experimentale et son Utilisation en Pedagogie.* Louvain : Nauwelaerts, pp. 366, FB. 260.

H. M. Bond (1959). *The Search for Talent.* Cambridge, Mass. : Harvard University Press ; London : Oxford University Press, pp. 57, 12s.

O. Hove (1958). *An Outline of Norwegian Education.* Oslo : Royal Norwegian Ministry of Foreign Affairs, pp. 189, 7s. 6d.

E. L. Trist and C. Sofer (1959). *Exploration in Group Relations.* Leicester University Press, pp. 68, 7s. 6d.

E. S. Ostrovsky (1959). *Father to the Child.* New York : Putnam, pp. xviii + 173, $3.75.

H. J. Eysenck (1959). *The Maudsley Personality Inventory.* University of London Press. Manual 2s. 6d., Keys 2s., Questionnaires 5d. each, 30s. 2d. per 100.

B. P. Wadia (1959). *The Building of the Home.* Bangalore, India : The Indian Institute of World Culture, pp. 56, Re. 1.00.

B. M. Bass and I. A. Berg (1959). *Objective Approaches to Personality Assessment.* New York and London : Van Nostrand, pp. x + 233, 37s. 6d.

N. H. Ledwith (1959). *Rorschach Responses of Elementary School Children : A Normative Study.* University of Pittsburgh Press, pp. xi + 185.

R. Thomson (1959). *The Psychology of Thinking.* Penguin Books, pp. 215, 3s. 6d.

N. E. Gronlund (1959). *Sociometry in the Classroom.* New York : Harper, xviii + 340, $4.50.

J. M. Morris (1959). *Reading in the Primary School.* National Foundation for Educational Research, Publication No. 12. Newnes, pp. xviii + 179, 15s.

J. Piaget and B. Inhelder, *et al.* (1959). *La Gènese des Structures Logiques Elémentaires.* Neuchatel : Delachaux and Niestle, pp. 295, Fr.s. 13.

E. Baranyai and E. Lenart (1959). *Az Irásbeli Közlés Gondolkodaslelektani Vonásai.* Budapest : Hungarian Academy of Sciences, pp. 315.

L. X. Magnifico (1950). *Education for the Exceptional Child.* New York and London : Longmans, Green, pp. xi + 371, 35s.

L. Cole (1959). *Psychology of Adolescence* (Fifth Edition.) New York : Rinehart, pp. xviii + 731, $7.00.

A. G. Hughes and E. H. Hughes (1959). *Learning and Teaching* (3rd Edition.) Longmans, Green, pp. x + 484.

K. C. Garrison and D. G. Force (1959). *The Psychology of Exceptional Children* (3rd Edition.) New York : Ronald Press, pp. vi + 586, $6.00.

F. Redl and W. W. Wattenberg (1959). *Mental Hygiene in Teaching* (2nd Edition.) New York : Harcourt, Brace, pp. xiv + 562, no price given.

J. Wittwer (1959). *Les Fonctions Grammaticales Chez L'Enfant.* Neuchatel : Delachaux and Niestle, pp. 295, Fr.s. 12.

W. G. Eliasberg (1959). *Psychotherapy and Society.* New York : Philosophical Library, pp. xviii + 223, $6.00.

J. Jacobi (1959). *Complex, Archetype, Symbol in the Psychology of C. G. Jung.* Routledge aud Kegan Paul, pp. xvii + 236, 21s.

C. W. Mills (1959). *The Sociological Imagination. New York* : Oxford University Press, pp. 234, 36s.

W. R. Jones (1959). *Bilingualism and Intelligence.* Cardiff : University of Wales Press, pp. 67, 4s. 6d.

BACK NUMBERS OF THE JOURNAL

The publishers, Methuen and Co., Ltd., would be glad to purchase issues of Vol. XV Part 3 (1945) and Vol. XVI Part 2 (1946), for 10s. each, provided they are in good condition.

CONFERENCE ON THE SCIENTIFIC STUDY OF MENTAL DEFICIENCY

JULY 24th TO 29TH, 1960

It is proposed, as a contribution to the observance of World Mental Health Year, that the American Association on Mental Deficiency and the Royal Medico-Psychological Association, with the co-operation of the National Association for Mental Health, should arrange in London a Conference dealing with the latest scientific and educational developments in mental deficiency, open to professional workers in this field.

The dates suggested have reference to the Meeting of the World Federation for Mental Health to be held in Edinburgh from August 7th to 12th, 1960, and plans would be made to allow participants to visit mental deficiency hospitals and training centres between the ending of the Conference and the opening of the Edinburgh Meeting.

The organisers of the Conference are anxious to hear as soon as possible from professional workers interested in the project.

Names of those interested should be sent to :

Dr. A. Shapiro,
c/o 39, Queen Anne Street,
London, W.1.

INDEX TO VOLUME XXIX, 1959

Articles :

PAGE

Academic Achievement, Two Personality Characteristics Related to : R. LYNN .. 213
Academic Achievement Tests for the Purpose of Academic Counselling at South African Universities, The Standardisation and Predictive Value of a Battery of: 136
Accomplishment Quotient Idea, An Historical and Critical Account of : A. R. CRANE 252
Assessments at Age Eleven, The Long-Term Follow-Up of : J. NISBET and J. BUCHAN 1
Bilingualism and Non-Verbal Intelligence : A Further Study of Test Results : D. G. LEWIS 17
Categories in Rating Scales : How Many and How Big ? : F. SANDON 155
Classification of Instances of Four Common Class Concepts by Children and Adults, The : M. ANNETT 223
Factor Analysis of the Wechsler Intelligence Scale for Children, A : A. E. MAXWELL 237
Factorial Analysis of the Evaluation of Scholastic Compositions in the Mother Tongue, A : C. REMONDINO 242
Factorial Study of Speed, Power and Related Temperament Variables, A : G. L. MANGAN 144
Function Fluctuation in a Mixed Group of Adolescents and Pre-Adolescents, The Extent and Factorial Structure of : C. C. ANDERSON 34
Graduate Student Teachers, A Follow-Up Study of some Former : M. COLLINS .. 187
Listening and Reading Comprehension for Boys and Girls of Primary School Age, An Experimental Investigation into the Relative Merits of : W. H. KING 42
Maladjusted Children in Grammar School : M. CHAZAN 198
Occupational Choice, Personality Factors in the Formation of : S. M. CHOWN .. 23
Piaget and Inhelder on the Child's Conception of Space, A Follow-Up Study of some Aspects of the Work of : K. LOVELL 104
Piaget's Schemata Concerning Children's Perception and Thinking, Experimental Examination of some of, and a Discussion of their Educational Significance : E. A. PEEL 89
Piaget's Theory of Sensori-Motor Development, The Behaviour of Idiots Interpreted by M. WOODWARD 60
Speech and Language Development of Five-Year-Old Children, The : O. C. SAMPSON 217
Symposium : The Development of Moral Values in Children :
  V.—The Influence of Religious Education upon the Development of Children's Moral Ideas : W. H. HILLIARD 50
  VI.—Punishment, Discipline and Educational Objectives : D. K. WHEELER .. 118
Teachers' Ratings of the Adjustment of their Pupils, The Value of : J. G. THORPE .. 207
Test Anxiety and the " Eleven Plus " Examination : I. SARNOFF, S. B. SARASON, F. F. LIGHTHALL, and K. S. DAVIDSON 9
Trends in Educational Psychology : S. WISEMAN 128

Comments, Criticisms, Correspondence :

Attainment Level on Leaving Certificate and Academic Performance at University (by E. M. Gould and J. G. M'Comisky) : R. R. DALE 74
On S. Wiseman's Article : " Trends in Educational Psychology " : E. C. HOPKINS, F. SANDON, S. WISEMAN 264
Studying the Meaning of Quantitative Terms : D. SHEPPARD 76
The Accomplishment Quotient : A Reply : C. BURT 259

Critical Notice :

Television and the Child (by H. T. Himmelweit, A. N. Oppenheim and P. Vince) : C. BURT.. 173

Research Notes :

R. BOWYER 162
A. J. M. FLOOK 261
H. MADDOX 72

Summaries of Research Theses : 77, 165, 267

Notes to Contributors : 283

Announcements : 280

Authors :

PAGE

Anderson, C. C. : The Extent and Factorial Structures of Function Fluctuation in a Mixed Group of Adolescents and Pre-Adolescents 34
Annett, M. : The Classification of Instances of Four Common Class Concepts by Children and Adults 223
Bowyer, R. : Research Note : The Importance of Sand in the World Technique : An Experiment 162
Burt, C. : Critical Notice : Television and the Child (by H. T. Himmelweit, A. N. Oppenheim and P. Vince) 173
Burt, C. : The Accomplishment Quotient : A Reply 259
Chazan, M. : Maladjusted Children in Grammar Schools 198
Chown, S. M. : Personality Factors in the Formation of Occupational Choice .. 23
Collins, M. : A Follow-Up Study of some Former Graduate Student Teachers .. 187
Crane, A. R. : An Historical and Critical Account of the Accomplishment Quotient Idea 252
Dale, R. R. : Critical Note : Attainment on Leaving Certificate and Academic Performance at University (by E. M. Gould and J. G. M'Comisky) 74
Flook, A. J. M. : Research Note : Note on the Use of New-Type Tests for Improving the Quality of Discussion in Discussion Groups 261
Hilliard, F. H. : Symposium : The Development of Moral Values in Children : V—The Influence of Religious Education upon the Development of Children's Moral Ideas 50
Hopkins, E. C. : Correspondence : On S. Wiseman's Article : 'Trends in Educational Psychology' 264
King, W. H. : An Experimental Investigation into the Relative Merits of Listening and Reading Comprehension for Boys and Girls of Primary School Age 42
Lewis, D. G. : Bilingualism and Non-Verbal Intelligence : A Further Study of Test Results 17
Lovell, K. : A Follow-Up Study of some Aspects of the Work of Piaget and Inhelder on the Child's Conception of Space 104
Lynn, R. : Two Personality Characteristics Related to Academic Achievement .. 213
Maddox, H. : Research Note : Mental Age Scales 72
Mangan, G. L. : A Factorial Study of Speed, Power and Related Temperament Variables 144
Maxwell, A. E. : A Factor Analysis of the Wechsler Intelligence Scale for Children 237
Nisbet, J., and Buchan, J. : The Long-Term Follow-Up of Assessments at Age Eleven 1
Peel, E. A. : Experimental Examination of some of Piaget's Schemata concerning Children's Perception and Thinking, and a Discussion of their Educational Significance 89
Remondino, C. : A Factorial Analysis of the Evaluation of Scholastic Compositions in the Mother Tongue 242
Sandon, F. : Categories in Rating Scales. How Many and How Big ? 155
Sandon, F. : Correspondence : On S. Wiseman's Article : 'Trends in Educational Psychology' 265
Sarnoff, I., Sarason, F. B., Lighthall, F. F., and Davidson, K. S. : Test Anxiety and the "Eleven Plus" Examination 9
Sheppard, D. : Critical Note : Studying the Meaning of Quantitative Terms .. 76
Sonnekus, M. C. H. : The Standardisation and Predictive Value of a Battery of Academic Achievement Tests for the Purpose of Academic Counselling at South African Universities 136
Thorpe, J. G. : The Value of Teachers' Ratings of the Adjustment of their Pupils .. 207
Wheeler, D. K. : Symposium : The Development of Moral Values in Children. VI.—Punishment, Discipline and Educational Objectives 118
Wiseman, S. : Trends in Educational Psychology 128
—— Reply to Correspondence 266
Woodward, M. : The Behaviour of Idiots Interpreted by Piaget's Theory of Sensori-Motor Development 60

NOTES TO CONTRIBUTORS

1.—The Editor is glad to consider articles in the fields of educational and child psychology, whether written from the theoretical, clinical, experimental or psychometric standpoints. It is, however, desirable that statistical matter (formulae, tables, discussions of techniques, etc.) should be kept to a minimum, so that the *Journal* will be of interest to statistically untrained readers. Such concepts as the mean, percentiles, standard deviation, correlation, chi-squared, standard error, critical ratio, significance and reliability, and general or group factors, can be employed in the text. But more complex concepts and methods should be explained, or confined to footnotes or an appendix, or published elsewhere.

Articles which have been, or are to be, published elsewhere cannot be accepted.

2.—*Length.* The usual range of length is from 2,500 to 4,500 words, and only very exceptionally will articles of more than 6,000 words be published. Short research reports of 500—1,000 words, such as outlines of theses, are welcomed. These should normally be submitted through the Head of the Department in which the research was carried out.

3.—*Arrangement.* Articles should be *preceded* by a summary or abstract of between 100 and 300 words, embodying the main conclusions. This should be typed directly after the title and author's name and place of work. A contents analysis, or list of headings, need *not* be included, but successive sections of the article should receive numbered headings, e.g. :

I. Introduction.
II. Outline of Previous Research.
III. Plan of the Investigation.
    (i) *Subjects and materials.*
    (ii) *Experimental Procedure.*
IV. Statement of Results.
V. Discussion and Conclusions.
VI. References.

4.—*Typing.* All papers must be typed (double or 1½-spaced) on one side of quarto or foolscap paper. There should be margins of at least one inch all round. Authors should consult articles in current numbers of the *Journal*, and copy their conventions in respect of capitalising the title and main section headings, italicising author's place of work and sub-section headings, indentation of paragraphs, the form of references and tables, etc.

5.—*Footnotes* should be few in number, and should be indicated in the text by using the signs *, †, ‡, §.

6.—*References* should normally be indicated by date following the author's name in the text, e.g., Spearman (1904), or (Spearman, 1904, p. 160), or—where there are two or more publications in the same year—Burt (1939a). However, it is occasionally more convenient to number the references throughout, and to type numbers in the text above the line, e.g., Spearman.[4] When there are two or three references only, these may be placed in footnotes in the text, instead of in a separate section at the end.

The list of references at the end should always be in alphabetical order of author's names and—for any one author—by order of date. The following standard form of references should be observed :

Dearborn, W. F., and Rothney, J. W. M. (1941). *Predicting the Child's Development.* Cambridge, Mass. : Sci.-Art.

(Note that Ltd., etc., after publisher's names, also the numbers of pages in the book, can be omitted.)
Burt, C. (1939a). The relations of educational abilities. *Brit. J. Educ. Psychol.*, **9, 55-71.**

(Note the absence of inverted commas and of capital initials in the titles of articles. The number of the volume is in arabic, not roman numerals. The part of the journal is not given, but page numbers are included in full).

The World List of abbreviations (as given in the December issues of *Psychological Abstracts*) is used for journals.

7.—*Tables and figures* are costly to print, and should be used only when they save space and/or are essential for effective presentation. The same material should not be presented in a table *and* a graph. Each table or figure should be numbered (in arabic figures), on a separate sheet, and its approximate position indicated in the typescript. Drawings should be in Indian ink on heavy unruled paper, the lettering on graphs being in pencil. Lines in tables should be ruled in pencil. Each table or graph should have a short title, sufficient for understanding without searching in the text.

8.—*Corrections.* The Editor cannot consider for publication articles which are seriously deficient in presentation. The typescript submitted (after preliminary consultation, if desired) should represent the final form in which the author wishes the paper to appear. Since changes in proof are costly, out of all proportion to the original setting, authors must be charged for any changes, insertions or deletions (other than printer's errors).

9.—*Reprints.* Authors receive twenty-five copies of their papers free. Extra copies, in multiples of twenty-five, may be had at cost price if the order is given when the proof is returned.

BUREAU OF EDNL. & PSYL. RESEARCH
LIBRARY

Vita Humana

Internationale Zeitschrift für Lebensaltersforschung
International Journal of Human Development
Journal International de Développement Humain

INDEX Vol. 2, No. 1 (1959)

H. Harmsen (Hamburg) : **Sozialhygienische Erhebungen zur Frage der Pensionierung.**
C. A. Dailey (Kansas City, Mo.) : **The Natural Structure of the Life History.**
R. J. Havighurst (Chicago, Ill.) : **Life Styles of Middle-Aged People.**
K. Haupt (Nürnberg) : **Formen der sozialen Eingliederung Vertriebener.**

INDEX Vol. 2, No. 2 (1959)

T. W. Moore (London) : **Studying the Growth of Personality. A Discussion of the Uses of Psychological Data Obtained in a Longitudinal Study of Child Development.**
A. Karsten (Helsinki) : **Adjustment to Old Age in Industry.**
K. Eyerth (Hamburg) : **Eine Untersuchung der Neger-Mischlingkinder in Westdeutschland.**

Editor : H. THOMAE, Erlangen

One number of 64 pages is published quarterly. One volume consists of 4 numbers and costs *sFr. 38.50* (postage incl.)

Within the Editorial Board scientists are represented working in the most important fields of research in human development within different countries. As an international journal " Vita Humana " publishes papers in English, French or German with summaries of the original papers in all three languages.

BASEL 11 (Switzerland) **S. KARGER** **NEW YORK**
Representative for Great Britain : INTERSCIENCE PUBL. LTD., London, W.2

OCCUPATIONAL PSYCHOLOGY

Editor : Alec Rodger

Volume 33, No. 3 JULY, 1959

Overtime and Financial Responsibility. By Leslie Buck and Sylvia Shimmin.

The Place of Personality Appraisal in Vocational Selection. By John Parry.

Attitudes of Union Leaders and Business Executives to Industrial Relations. By Kenneth F. Walker.

A Sociological Case Study of Operator Training : Part I. By Shirley Wilson.

Professor Landsberger on Hawthorne. By Peter Cavanagh.

Book Reviews.

Other Books Received.

Annual Subscription 40 *shillings.*

National Institute of Industrial Psychology
14, Welbeck Street, London, W.1.

BUREAU OF EDNL. & [illegible] RESEARCH
LIBRARY